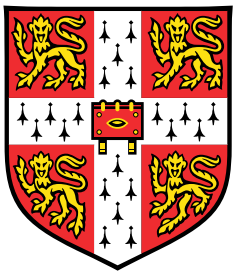
Feedback for neuronal system identification



Thiago Bassinello Burghi

Supervisor: Prof. R. Sepulchre

Department of Engineering
University of Cambridge

This thesis is submitted for the degree of
Doctor of Philosophy

Churchill College October 2019

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Introduction and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Introduction and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Introduction and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

Thiago Bassinello Burghi
October 2019

Abstract

Feedback for neuronal system identification
Thiago Bassinello Burghi

In order to estimate reliable models from noisy input-output data, system identification techniques usually require that the data be generated by a process with a fading memory. Non-equilibrium systems such as neuronal and chaotic models lack a fading memory. Their identification is challenging, in particular in the presence of input noise. In this thesis, we propose a methodology based on the prediction-error method for the identification of neuronal systems subject to input-additive noise. We build on the fundamental observation that while a neuronal model does not have a fading memory, it can be transformed into a fading memory system by output feedback. Our ideas can be generalized to any non-equilibrium system sharing this property.

At the core of the methodology is the use of output feedback in experiment design. We provide a theoretical justification for this design choice, which has been exploited in neurophysiology since the invention of the voltage-clamp experiment. To investigate the problem of feedback for identification, we first address the estimation of simple non-equilibrium systems in Lure form, and show that feedback allows estimating the nonlinearity in a static experiment. We then address the estimation of conductance-based models. Assuming that an informed choice can be made on the elements of the model structure, we show that consistent parameter estimates can be obtained when noise is only present at the system input. Finally, we approach the problem from a black-box perspective, and propose identifying the neuronal internal dynamics using a universal approximator with Generalized Orthogonal Basis Functions.

To my parents, Cláudia and Nilton

Acknowledgements

I would like to thank my supervisor, Rodolphe Sepulchre, for encouraging me to dive into all the interesting topics I explored during my PhD. His patience and support were crucial for the completion of this thesis.

I would also like to thank Maarten Schoukens for the fruitful collaboration and for helping me to navigate the world of nonlinear system identification.

This was a long journey I started with Ilario Cirillo, Luka Ribar, Tomas van Pottelbergh and Alessandro Morelli, and it would certainly have been harder to arrive here without their companionship.

The Control Group at Cambridge is a great place for discussing ideas, and I thank everyone there for the interesting exchanges, in particular Christian Grussler, Fulvio Forni and Tim O'Leary.

I also acknowledge the financial support I received from the Brazilian government through its higher education agency CAPES. Without their support, studying at Cambridge would have only been a dream.

I thank my parents Cláudia and Nilton for all the support they have given me throughout my education. Thanks to my sister Bianca for planning some much needed vacations.

Thanks to Fabian for being a great flatmate and an even better friend.

Last but not least, I am immensely lucky to have met Sacha. Her love and company during the past few years have helped me keep going during the hard times.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

One of the most influential nonlinear system identification experiments ever devised was carried out by the neurophysiologists Hodgkin and Huxley in the late 1940s to investigate the mechanisms behind the generation of action potentials in neuronal axons. Their experiments relied on a simple, but powerful idea: to use feedback to turn a difficult identification problem into a tractable one. As empirically verified by Hodgkin, Huxley, and the thousands of neurophysiologists who still use their experimental protocol to this day, neuronal dynamics can be stabilized by output feedback. This thesis explores the theoretical implications of this property in the context of nonlinear system identification.

System identification

In some respects, nonlinear system identification is a mature field. Our understanding of nonlinear identification problems has advanced to a point where identifying a wide class of *mildly* nonlinear systems has become routine. The fundamental property shared by models in that class is a fading memory: systems with a fading memory forget initial conditions, as well as the distant past of their inputs. Thanks to a solid body of existing research, one can use an increasing number of methods to identify such systems using Volterra models, Wiener/Hammerstein models, nonlinear auto-regressive (NARX) models, and artificial neural networks. Provided the fading memory condition (or a number of closely related properties) holds, such methods provide theoretical guarantees that the identification procedure will yield satisfactory results in the face of noise and model uncertainties.

In contrast, research concerning the identification of systems lacking a fading memory – *severely* nonlinear systems, in the terminology of [10] – is still in its infancy.

It was not until very recently that the system identification community turned their attention to the identification of severely nonlinear biological systems which, due to their excitable nature, lack the fading memory property. This work focuses on a prime example of such a system: the biophysically realistic neuronal model.

Neuronal systems

A neuron is a cell which is capable of actively changing the electric potential across its membrane in response to external stimuli. Those changes in potential are induced by the flow of ionic currents through proteins embedded in the membrane, called ion channels. Depending on the input stimulus and on the properties of their ion channels, neurons are capable of displaying nonlinear behaviours such as action potentials (also known as spikes), as well as oscillations that can be periodic, random, or chaotic in nature.

Neurons are excitable, meaning that they have a resting state, and can discharge an action potential – or a *burst* of action potentials – whenever their input exceeds a certain threshold. Although it is simple to explain excitability and the threshold at an intuitive level, these concepts are difficult to rigorously define. Part of the reason for this difficulty lies in the lack of biomolecular first principles that allow us to model neuronal ion channels [56, 74, 5].

Neuronal models are the fundamental tool for doing research in computational neuroscience, and the fact that neurons vary in their dynamical properties across brain regions and across individuals makes it necessary to have system identification tools that can deliver reliable models based on data that is simple to obtain.

Feedback-invariant properties of neuronal systems

Hodgkin and Huxley's pioneering work established the paradigm of conductance-based modeling, which is still used in current computational neuroscience studies. Their model combined a phenomenological explanation for the inner workings of ion channels with two key system properties grounded in physical reality. In the language of control theory, the two key properties are invariant under output feedback: neuronal systems have a relative degree of one, and their inverse dynamics has a fading memory. In the language of neurophysiology, those properties correspond, respectively, to the fact that the membrane acts as a capacitor, relating the temporal change of voltage to the sum of ion currents, and to the fact that ion currents only depend on the recent past of the membrane voltage.

Thesis outline and contributions

The main question posed in this thesis is *how to identify systems that possess the feedback-invariant properties of neuronal systems.* In particular, we will focus on using the prediction error method (PEM).

The unifying thread in this thesis is the use of output feedback as a means to approach the question above. In the next chapter, we formulate the problem in detail, discussing the limitations of neuronal system identification, and reviewing the literature to understand how the problem has been approached by other researchers. While all the results presented in Chapter 2 are known, the discussion of neuronal state-space models in Section 2.5.2 is original to this thesis.

We close this introduction with a brief summary of the main contributions of the thesis, which are divided into three main chapters.

Feedback identification of Lure systems

In Chapter 3, we rely on output feedback to develop a closed-loop identification method for excitable and chaotic Lure systems. In this method, we first identify the system nonlinearity in a static identification stage; then, knowledge about the nonlinearity can be used to turn the problem into that of identifying a mildly nonlinear system.

This contribution has been published jointly with Maarten Schoukens and Rodolphe Sepulchre in [15].

Feedback identification of conductance-based models

In Chapter 4, we extend the simple idea of Chapter 3 so as to develop a procedure for consistent identification of conductance-based models under input-additive noise using the PEM method. This chapter provides a theoretical justification for methods in which a neuronal model is estimated by including pre-identified ion channel kinetics in the model structure. In addition, in Chapter 4 we show that output feedback can be used to estimate approximate eigenvalues of a neuronal system's linearized internal dynamics. This can be used to determine a model structure for the nonlinear identification problem, and to determine the voltage-dependent time constants of a conductance-based model.

Parts of this contribution have been accepted for publication [16] jointly with Maarten Schoukens and Rodolphe Sepulchre.

Identification of neuronal systems

In Chapter 5, we contribute to the question in two ways: first, we explore the remarkable fact that a noisy input can sometimes induce a fading memory, dispensing with the need for feedback. We then propose using a black-box model structure containing Generalized Orthonormal Basis Functions (GOBFs) to identify the neuronal inverse dynamics. The GOBFs poles can be chosen in a way that facilitates the identification of the localized behavior of the system in the vicinity of a bifurcation. The methods of Chapter 4 can be used to approximately identify those poles.

Parts of this contribution have been accepted for publication [17] jointly with Maarten Schoukens and Rodolphe Sepulchre.

1.1 Notation

In this thesis, we use the notation $\mathbb{R}_+ = [0, \infty)$ and $\mathbb{Z}_+ = \{0, 1, 2 \dots\}$. We use subscripts to denote the value of a scalar or vector-valued discrete-time signal at a particular moment of time. The number 0 is treated as a scalar or as a vector, with the dimension implied by the context in which it is used.

The norms $\| \cdot \|_p$ denote the usual p-norms of $\mathbb{R}^n$. We denote by $\ell_\infty^n(\mathbb{Z}_+)$ the space of all $\mathbb{R}^n$-valued sequences u_k defined on $\mathbb{Z}_+$ such that $\|u\|_{\ell_\infty} < \infty$, where $\|u\|_{\ell_\infty} = \sup_{k \geq 0} \|u_k\|_\infty$ (we drop n from the notation whenever $n = 1$). The set $\mathbb{R}(\mathbb{Z}_+)$ is the set of real-valued sequences defined on $\mathbb{Z}_+$. Similarly, $\mathcal{L}_\infty(\mathbb{R}_+)$ is the set of all real-valued bounded functions defined on $\mathbb{R}_+$. The operator $\sigma_{\max}[\,\cdot\,]$ denotes the largest singular value of a matrix.

For a discrete-time variable x_k, $k \geq 0$, the signal up to time k is denoted by $x_{[0,k]} = (x_k, x_{k-1}, \dots, x_0)$.

Chapter 2

Problem formulation and state-of-the-art

2.1 Problem formulation

The main purpose of this thesis is to address the problem of identifying nonlinear input-output systems that have the feedback representation shown in Figure 2.1. The external input is denoted by $i(t)$ (the conventional symbol for electric current), and the output is denoted by $v(t)$ (the conventional symbol for voltage). The feedback loop is composed of a linear time-invariant (LTI) passive system in the forward path, and a nonlinear fading memory system in the feedback path.

In this section, we motivate this problem and highlight some important questions that we aim to answer in this thesis. In Sections 2.2-2.5, we review and provide detailed definitions of the main concepts which are discussed here.

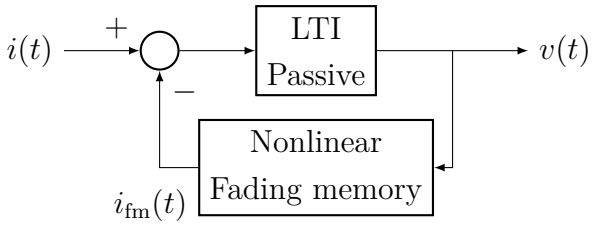


Fig. 2.1 A nonlinear feedback circuit.

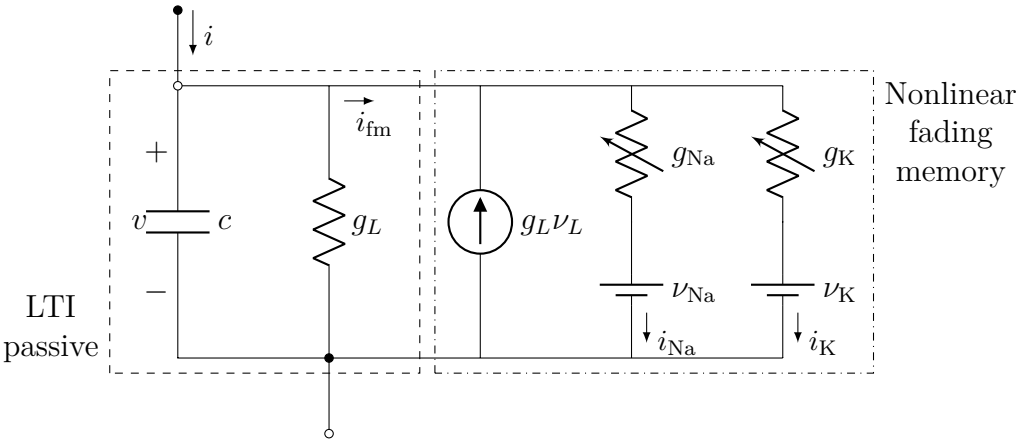


Fig. 2.2 The Hodgkin-Huxley circuit, and its relationship to the feedback structure of Figure 2.1. The currents i_{Na} and i_{K} are the sodium and potassium ion currents, respectively. In addition, the model possesses a leak current $i_{\mathrm{L}} = g_L(v - \nu_L)$. The leak conductance g_L is static, while g_{Na} and g_{K} are voltage-dependent adaptive conductances which can be modeled by an operator with a fading memory (this point is further developed in Chapters 2 and 4). The model is detailed in Section 4.3.1.

2.1.1 Neuronal models motivate the feedback structure

The notation we have chosen for the input and the output variables reflects the fact that we are motivated primarily by the problem of identifying neuronal circuits. In this context, the input represents an applied external current $i(t)$, and the output represents a neuronal membrane voltage $v(t)$. The passive element models the neuronal membrane dynamics, while the element with a fading memory models the total sum of ionic currents flowing through the membrane. In particular, this feedback architecture is central to neuronal conductance-based modeling, the framework introduced by Hodgkin and Huxley in their seminal work [57]. The relationship between the elements of the feedback structure of Figure 2.1 and the components of the electrical circuit representing Hodgkin and Huxley's model is shown in Figure 2.2.

The feedback structure of Figure 2.1 is very general: it encompasses not only single-compartment neuronal models, such as the Hodgkin-Huxley model, but also multi-compartment models and neuronal network models. The difference between these types of models is related to the dimension of the variables $i(t)$ and $v(t)$. In single-compartment models, $i(t)$ and $v(t)$ are scalar signals, and the system is said to be single-input-single-output (SISO). In multi-compartment and network models, $i(t)$ and $v(t)$ are vector-valued signals, and the system is said to be multiple-input-multiple-output (MIMO). In the MIMO case, inter-compartmental resistive currents

are included in the passive element of the interconnection, while synaptic currents contained in the network are included in the fading memory element.

Although there is no conceptual difference between the MIMO and the SISO case, this thesis will focus on the problem of identifying SISO models with the structure of Figure 2.1.

2.1.2 Open-loop and closed-loop behaviors

The feedback structure of Figure 2.1 accounts for the rich behavior[1] of neuronal systems. While a fading memory system is characterized by an equilibrium behavior for constant input signals, it is the feedback interconnection of fading memory elements that accounts for excitability and the non-equilibrium behavior of a neuronal model.

Non-equilibrium behaviors include relaxation oscillations, bistability, bursting, and chaos. In contrast, each block of the feedback structure has a much simpler open-loop behavior. A passive system can only dissipate energy, and passivity implies stability of a linear system [134, 67]. The fading memory property implies properties typical of stable linear systems, such as unique steady-states [13] and entrainment by periodic inputs [121].

The stark difference between the behavior of the closed-loop feedback system and the open-loop behavior of its constituent components can be illustrated with input signals that are commonly used in system identification. In Figures 2.3 and 2.4, we contrast the behavior of the Hodgkin-Huxley model of Figure 2.2 with the behavior of its fading memory component (the element modeling the sum of ion currents).

The step responses of the element with a fading memory, shown on the left of Figure 2.3, are reminiscent of the step responses of a linear system. The same is true of the responses of that element to white noise, shown on the left of Figure 2.4.

In contrast, the step responses of the closed-loop feedback system, shown on the right of Figure 2.3, are "severely nonlinear"; in particular, the system has a localized ultra-sensitivity to the input. Doubling the input step amplitude from $2\,\mu A/cm^2$ to $4\,\mu A/cm^2$, for instance, results in an increase in the peak amplitude of the output signal of almost two orders of magnitude. In neurophysiological terms, the step of $2\,\mu A/cm^2$ results in a small amplitude *graded potential*, while the step of $4\,\mu A/cm^2$ results in a large amplitude *action potential* – a spike.

[1] The word *behavior* has a precise meaning under the framework of Willems' behavioral theory [150]. The (manifest) behavior of an open system (a system with inputs and outputs) is the set of input-output signal trajectories that are allowed by the system.

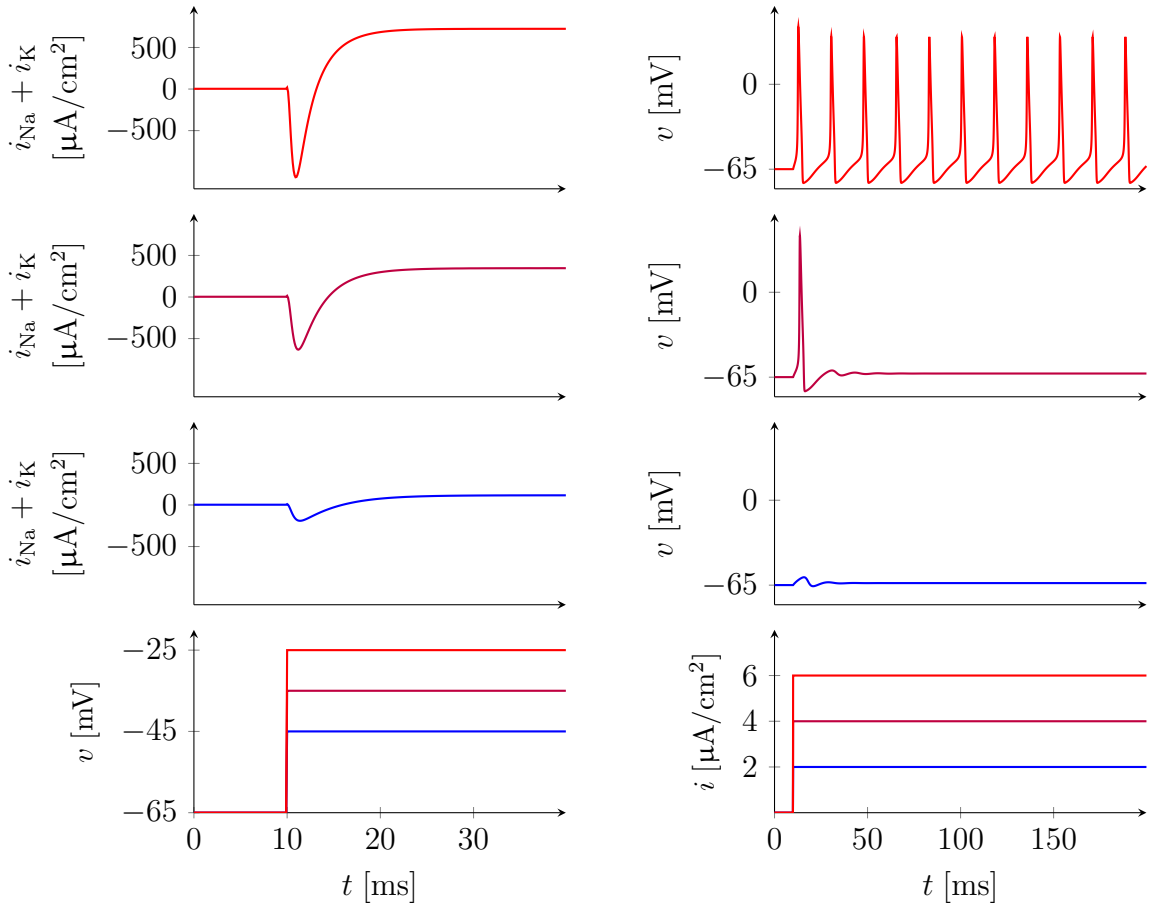


Fig. 2.3 Left: Open-loop behavior of the fading memory element of the Hodgkin-Huxley model subject to steps of different amplitudes at the input. Right: Behavior of the Hodgkin-Huxley model subject to steps of different amplitudes at the input.

The localized ultra-sensitivity of neuronal models results in a sensitive dependence on the realization of a stochastic input, as shown on the right of Figure 2.4. When the system is subject to a stochastic signal such as white noise, the occurrence of an output spike (or of a sequence of spikes) is a stochastic event that depends on the realization of the input. Changing the realization of the input may drastically change the precise location of spikes along the time axis.

2.1.3 Neuronal behaviors are sensitive to parameter changes

Neuronal models can display high sensitivity to changes in their open-loop behavior and the way their fading memory element (which models the total ion current) is parameterized. This is illustrated through Figures 2.5 and 2.6, where we compare the

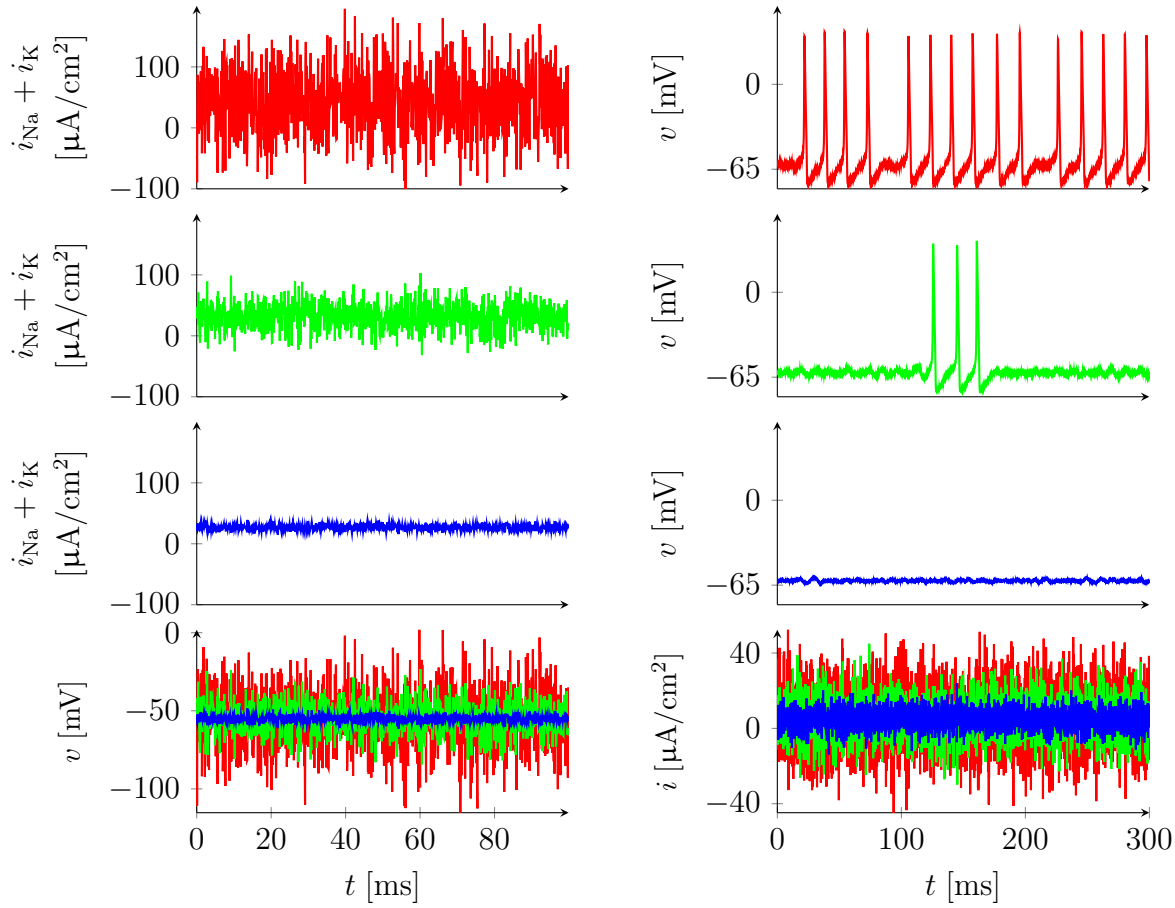


Fig. 2.4 Left: Open-loop behavior of the fading memory element of the Hodgkin-Huxley model subject to white noise of different variances at the input. Right: Behavior of the Hodgkin-Huxley model subject to white noise of different variances at the input.

open-loop and closed-loop behaviors of two different models with similar parameter values.

In Figure 2.5, we show that the two open-loop fading memory elements produce outputs which are "close" when subjected to the same input: computing the normalized root-mean-square error (NRMSE) between the two outputs, we obtain a value of 0.9041 (close to the maximum value of 1). However, Figure 2.6 shows that the closed-loop behaviors of the two models are qualitatively and quantitatively different from each other. Applying the same input current to both models yields two rather different spiking traces.

This small example shows part of the challenge in the neuronal system identification problem: similar models may produce different excitable behaviors. Later on, we will see that similar excitable behaviors can only arise if the fading memory elements are

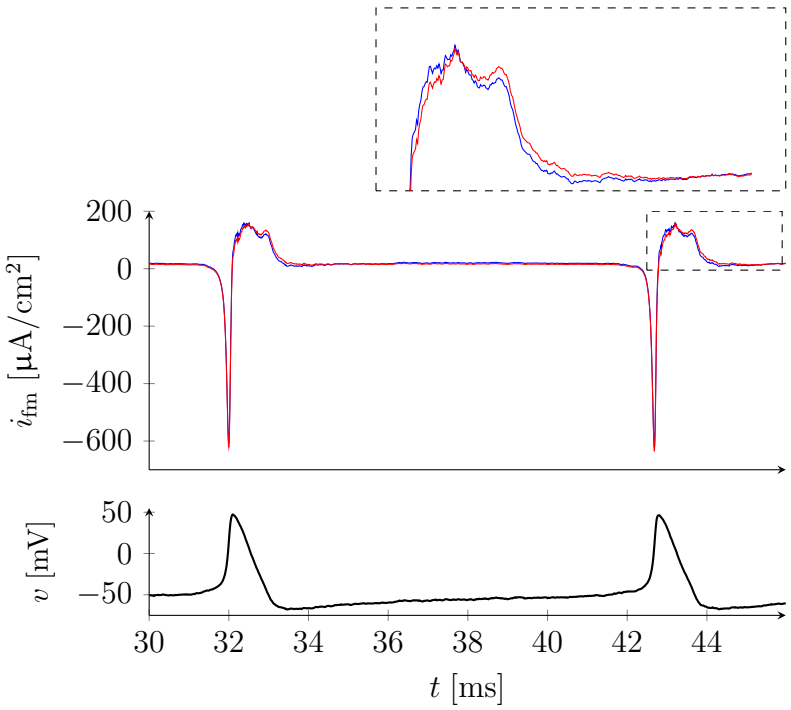


Fig. 2.5 Comparison of open-loop behavior of the fading memory element of two slightly different models. When the fading memory elements are subject to the same voltage input v (bottom panel), their outputs i_{fm} (blue and red traces in the top two panels) remain reasonably close to each other (NRMSE = 0.9041).

close *locally*, and that the problem of estimating good neuronal models can be directly related to a classical problem of closed-loop identification.

2.1.4 Neurons are noisy

In this thesis, we will often consider the noise setup shown in Figure 2.7: the known excitation signal $i(t)$ is corrupted by an unmeasured input-additive noise term $e(t)$. This noise setup is again motivated by neuronal applications. Neurons are noisy systems, and noise can affect those systems in many ways. Noise in neurons can be classified as intrinsic or external [46]. In the context of neuronal identification, the input-additive noise shown in Figure 2.7 could be used to model the combination of external and intrinsic membrane current noise. The external noise could correspond to the summation of many filtered synaptic currents, coming from the neuron's dendrites, in an experiment performed *in vivo*. The intrinsic noise could come from unmodeled stochastic effects in the neuron's ion channels. The input-additive noise framework

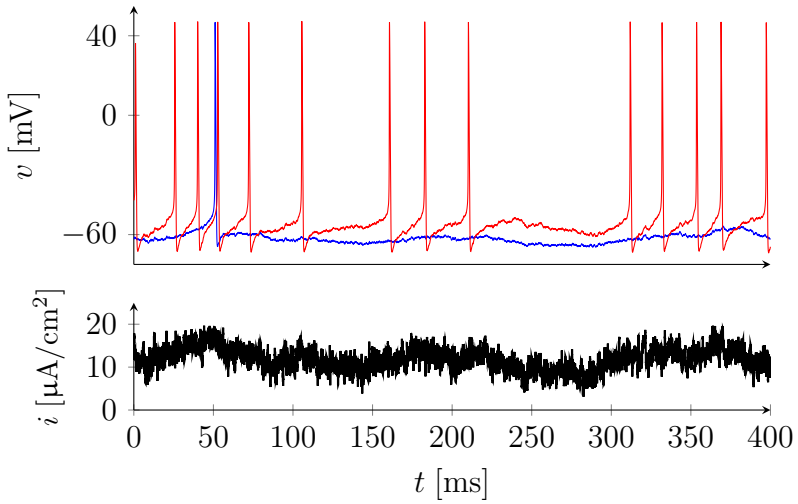


Fig. 2.6 Comparison of the closed-loop behavior of the systems with the open-loop fading memory behavior shown in Figure 2.5. Despite having "close" open-loop behaviors, indicating closeness of model parameter values, the closed-loop behaviors of the models are substantially different from each other.

is a simple representation of those effects that allows us to analyze the identification problem.

The input-additive noise models a phenomenon long known to experimentalists: when the same known excitation $i(t)$ is applied to a neuron in a sequence of experiments, the neuron's response changes from one experiment to the other (Figure 2.8). This is hardly any news for experts in system identification. In fact, the central question of classical linear system identification theory [77] is how to deal with the fact that the collected data contains but one of many possible realizations of the noise. This question is critical to the identification of neuronal systems, as the hypersensitivity to the input may cause different noise realizations to result in drastically different outputs.

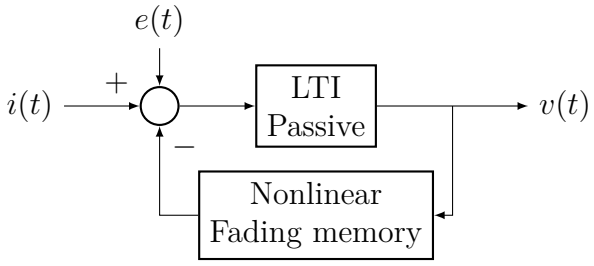


Fig. 2.7 A nonlinear feedback circuit with input-additive noise.

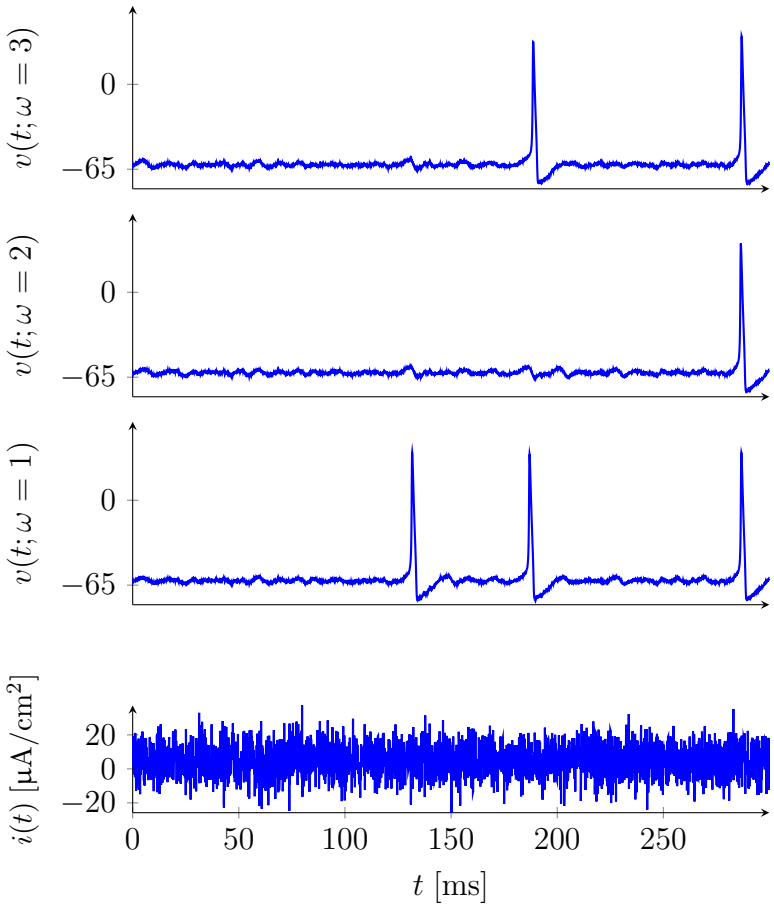


Fig. 2.8 Three different voltage outputs of the Hodgkin-Huxley model subject to the same realization of Gaussian white noise applied current $i(t)$ but three different realizations of Gaussian input-additive white noise: $e(t;\omega = 1)$, $e(t;\omega = 2)$ and $e(t;\omega = 3)$. The applied current has a mean of 5 [μA/cm^2] and a standard deviation of 9 [μA/cm^2], while the input-additive noise has zero mean and a standard deviation of 2 [μA/cm^2]. Voltages are in [mV], and currents are in [μA/cm^2].

2.1.5 Simulation and prediction

When noise is considered in the identification problem, it is common to distinguish the problem of *prediction* from that of *simulation*. The problem of prediction involves finding a model (called a predictor model) that makes use of past values of the input and the output of a ground-truth system — possibly disturbed by noise — so as to predict the current (or future) value of the system's output. The problem of simulation involves finding a model (called a simulation model) that can reproduce the undisturbed output of the ground-truth system in case both the model and the system are subject to the same input [77].

In view of that distinction, it is useful to clarify what we mean by identifying the closed-loop feedback system of Figure 2.7 under the input-additive noise assumption. We will not attempt to solve the problem of prediction. We will rather seek a model that can reproduce the output $v(t)$ of the system in case both the system and the model are subject to the same input $i(t)$ and *the same realization* of the noise $e(t)$. In other words, we seek a model that can reproduce the deterministic behavior of the ground-truth system. To find such a model, however, we still have to deal with the fact that the data is corrupted by the stochastic signal $e(t)$, which is not measured.

2.1.6 Voltage-clamp vs. current clamp

In neurophysiology, an experiment in which a neuron is driven directly by a pre-established current $i(t)$ is called a "current-clamp" experiment. In contrast, an experiment in which the voltage-gated ion currents (lumped in the fading memory element) are driven by a pre-established voltage $v(t)$ is called a "voltage-clamp" experiment.

The voltage-clamp experiment in the sense just described is of course idealized. In practice, the voltage-clamp experiment is made possible by the use of output feedback, and it is a voltage reference signal that is pre-established by the experimenter. In Chapter 4, we show that the output feedback in a voltage-clamp experiment can endow a neuron with a fading memory. This allows us to connect this classical experiment of neurophysiology to rigorous system identification methods.

2.2 Fading memory

In this section, we briefly introduce topics of nonlinear system theory which will be used throughout the thesis. We first introduce the fading memory property. It is an *input-output* system property, that is, it is a property of the system regarded as an operator between function spaces. We will see, later on, that fading memory is central to nonlinear system identification. We then introduce the concept of exponentially contracting dynamics. This is an *internal* system property, that is, it is a property on the system state-space representation. It can be thought of as an analog (albeit stronger) version of fading memory for state-space systems.

2.2.1 Definitions

The notion of fading memory was proposed by Boyd and Chua in [13]. The authors originally defined this property for continuous-time and discrete-time systems with

inputs and outputs defined on $\mathbb{R}$ and on $\mathbb{Z}$, respectively. In this section, we introduce the key concepts in the discrete-time setting, for systems with inputs and outputs defined on $\mathbb{Z}_+$.

A system operator $F : \ell_\infty^n(\mathbb{Z}_+) \to \mathbb{R}(\mathbb{Z}_+)$ is said to be *causal* if, for any input $u \in \ell_\infty^n(\mathbb{Z}_+)$, $(Fu)_k$ does not depend on u_n for $n > k$. It is said to be *time-invariant* if, letting $v_k = u_{k-n}$, we have $(Fv)_k = (Fu)_{k-n}$ for all $k \geq n$, and $(Fu)_k = 0$ for $k < n$ whenever $u_k = 0$ for $k < n$.

Definition 1 ([107])**.** Let $\mathcal{U}$ be a subset of $\ell_\infty^n(\mathbb{Z}_+)$. A discrete-time system operator $F : \ell_\infty^n(\mathbb{Z}_+) \to \mathbb{R}(\mathbb{Z}_+)$ is said to have fading memory on $\mathcal{U}$ if there exists a decreasing sequence $w : \mathbb{Z}_+ \to (0, 1]$ with $\lim_{k\to\infty} w_k = 0$ such that given an $\epsilon > 0$, there is a $\delta > 0$ such that for every $u, \tilde{u} \in \mathcal{U}$ and every $k \in \mathbb{Z}_+$,

$$\max_{m\in\{0,1,\ldots,k\}} \|u_m - \tilde{u}_m\|_\infty w_{k-m} < \delta \implies |(Fu)_k - (F\tilde{u})_k| < \epsilon \tag{2.1}$$

It can be shown that a time-invariant system with fading memory is causal [13, 107]. All systems considered in this thesis are causal and time-invariant, but many of them do not have a fading memory. Intuitively, a fading memory means that two inputs which are close in the recent past, $m \leq k$, result in outputs which are close at the present time k; the memory of the system regarding the distant past of those inputs, $m \ll k$, fades away.

Boyd and Chua showed that systems with a fading memory can be uniformly approximated on sets of uniformly bounded inputs by two important classes of *black box* models: finite Volterra series and time-delay neural networks. More precisely, their outputs converge in the uniform norm, as the number of terms in the approximating structure increases. This fact has important implications for system identification, as it provides a justification for the use of such nonlinear black-box model structures.

A slightly different notion of fading memory, called approximately-finite memory, was later proposed by Sandberg [120, 107]. This notion was also introduced for continuous-time and discrete-time systems, with the latter being defined as follows:

Definition 2 ([107])**.** Let $\mathcal{U}$ be a subset of $\ell_\infty^n(\mathbb{Z}_+)$. A discrete-time system operator $F : \ell_\infty^n(\mathbb{Z}_+) \to \mathbb{R}(\mathbb{Z}_+)$ is said to have approximately-finite memory on $\mathcal{U}$, denoted $F \in \mathcal{A}(\mathcal{U})$, if for any given $\epsilon > 0$, there is a $\Delta \in \mathbb{N}$ such that

$$|(Fu)_k - (FW_{k,\Delta}u)_k| < \epsilon, \quad k \in \mathbb{Z}_+ \tag{2.2}$$

for all $u \in \mathcal{U}$, where $W_{k,\Delta}u$ is a windowed version of u, given by

$$(W_{k,\Delta}u)_\kappa = \begin{cases} u_\kappa, & k-\Delta \leq \kappa \leq k \\ 0, & \text{otherwise} \end{cases} \tag{2.3}$$

Sandberg showed that systems in $\mathcal{A}(\mathcal{U})$ share the approximation properties of systems with fading memory [120], provided they satisfy a certain continuity condition. It turns out this continuity condition accounts for the difference between the concepts of fading memory and approximately-finite memory:

Proposition 1 ([107])**.** *Let $\mathcal{U}$ be any closed ball in $\ell^n_\infty(\mathbb{Z}_+)$ centered at the origin. Then $F : \ell^n_\infty(\mathbb{Z}_+) \to \mathbb{R}(\mathbb{Z}_+)$ has fading memory on $\mathcal{U}$ if and only if $F \in \mathcal{A}(\mathcal{U})$ and $(F\,\cdot)_k : \ell^n_\infty(\mathbb{Z}_+) \to \mathbb{R}$ is continuous on $\mathcal{U}$ for each $k \in \mathbb{N}$.*

Dealing with the notion of approximately-finite memory allowed Sandberg to demonstrate the connection between fading memory and the circle criterion, an important result in control theory used to check the absolute stability[2] of nonlinear Lure systems — the feedback interconnection of a LTI system with a static nonlinearity. In particular, a test similar to the circle criterion can be used to check whether a nonlinear Lure system possesses an approximately-finite memory [122, 22]. Such a test can be useful for analysis, since, for some nonlinear systems, it is often hard to verify the condition (2.1) or the condition (2.2) directly.

To illustrate the approximately-finite memory property, consider the input set $\overline{\mathcal{U}}$ given by all sequences $u : \mathbb{Z}_+ \to \mathbb{R}$ such that $\|u\|_{\ell_\infty} \leq \beta$, for some large radius β. We show in Figures 2.9 and 2.10 the behaviors of two different SISO causal and time-invariant systems: one is in $\mathcal{A}(\overline{\mathcal{U}})$, while the other is not.

The system whose behavior is shown in Figure 2.9 is in $\mathcal{A}(\overline{\mathcal{U}})$: we can find a sufficiently long time window Δ in order to achieve a small tolerance on the closeness of the outputs of the system. In Figure 2.10, we show the behavior of a simple oscillator (a discretized version of the FitzHugh-Nagumo model, introduced in Chapter 3) that is perturbed by a transient input. The system is not in $\mathcal{A}(\overline{\mathcal{U}})$, as any bounded transient input will have a permanent impact on the output of the system in the form of a constant phase difference. The phase difference ensures that there is a lower bound b on how close the outputs can be, and we cannot find a finite Δ that satisfies (2.2). Oscillators, bistable switches and chaotic systems with inputs and outputs are all examples of systems that, for a sufficiently large radius β, do not have a fading memory on $\overline{\mathcal{U}}$.

[2]For a definition of absolute stability, see e.g. [67, Chapter 7].

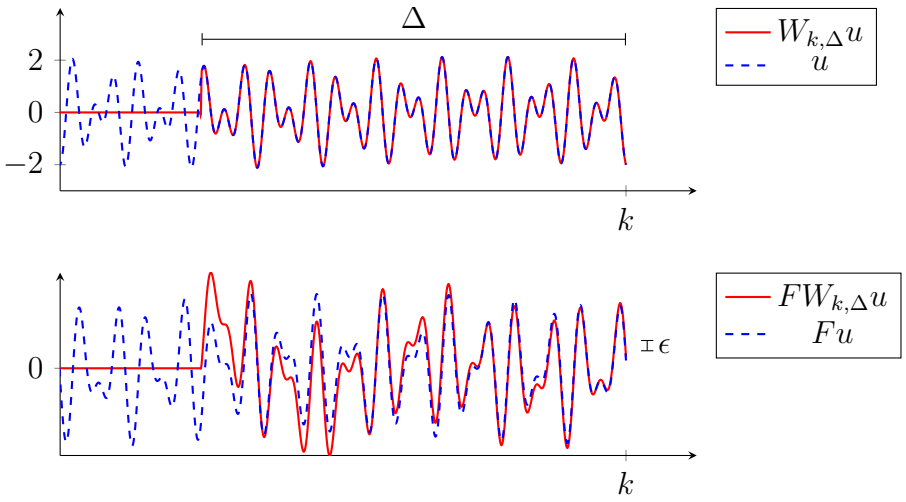


Fig. 2.9 Behavior of a causal and time-invariant system with approximately-finite memory (and fading memory) on the closed ball $\overline{\mathcal{U}}$. This particular system is a zero-order hold discretization of the linear stable system $G(s) = (s+1)/(s^2+0.8s+4.16)$.

We will use the concepts introduced above in Chapter 3 in order to tackle the identification problem of a simplified version of the interconnection of Figure 2.1. We will also use it in Chapter 5, where we use a model structure that is based on the approximation properties of fading memory systems.

2.2.2 Exponential contraction

If a fading memory input-output operator can be realized by a state-space model, then the memory encoded in the initial state variables of the system must fade away as time progresses. This behavior, which had already been noted by Boyd and Chua in their original paper [13], is typical of a class of state-space systems which have an exponentially contracting dynamics [81, 149, 8]. The key property of exponentially contracting systems is that trajectories which are infinitesimally close to each other converge to each other exponentially fast. It turns out that if this property holds on the whole state-space, many *global* properties are implied.

While we defer a rigorous treatment of this subject and its connections to neuronal system identification to Chapter 4, we introduce below a fact that illustrates the power of contraction theory for the global analysis of dynamical systems:

Proposition 2 ([8])**.** *Consider the autonomous system $\dot{x} = f(x)$, with f a smooth function. If the system is exponentially contracting on the whole state-space, there*

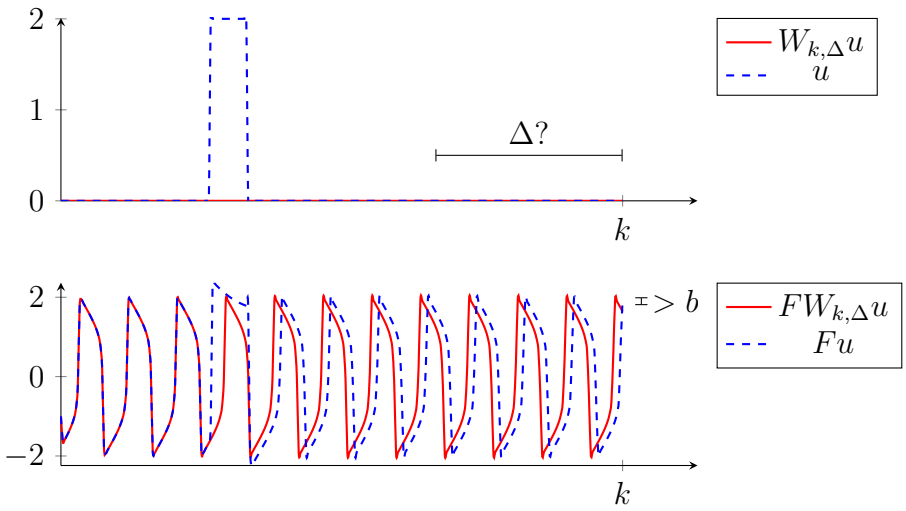


Fig. 2.10 Behavior of a causal and time-invariant system without fading memory on the closed ball $\overline{\mathcal{U}}$. The transient pulse permanently affects the output of the system: a constant phase difference exists between the original output and the output due to a windowed input which does not contain the pulse. This particular system is a forward-Euler discretization of the Fitzhugh-Nagumo model described on page 55.

exists a unique finite equilibrium, and all trajectories converge to this equilibrium exponentially fast.

In Chapter 4 (Section 4.2.1), we provide an original contribution that rigorously links the discrete-time versions of the fading memory property and the exponential contraction property, and show that systems with an exponentially contracting dynamics have a strong form of fading memory: their memory fades exponentially fast. In Chapter 4, contraction theory will aid us in tackling the identification problem of the interconnection of Figure 2.1 using a state-space framework.

2.3 Nonlinear system identification

In this section, we introduce and review some recent advances in nonlinear system identification, focusing on the parametric prediction error method (PEM) and on the importance of the fading memory property to system identification.

We argue that the fading memory property is central to three problems in parametric system identification: the asymptotic analysis (convergence and consistency) of estimates obtained in a stochastic setting; the search for a solution that minimizes the identification criterion function; and the approximation of nonlinear systems by linear ones – with the purpose of finding initial parameter estimates for nonlinear estimation.

2.3.1 Prediction-error method

In this section, we briefly present the classical prediction-error method (PEM), and review some recent advances related to that method in the context of nonlinear system identification.

The basic framework

In the classical (or "frequentist") framework of the prediction-error method [76, 77], we start from the assumption that there exists a ground truth system to be identified. Suppose the ground truth is a nonlinear stochastic discrete-time system represented by

$$y_k = F_k\left(u_{[0,k]}, e_{[0,k]}; x_0\right), \tag{2.4}$$

where $u_k \in \mathbb{R}^{n_u}$ is the system's input, $y_k \in \mathbb{R}^{n_y}$ is the system's output, the $e_k \in \mathbb{R}^{n_y}$ are random variables such that $\mathbb{E}[e_k \,|\, e_{[0,k-1]}] = 0$, the $F_k(\cdot, \cdot; x_0)$ form a sequence of deterministic mappings, and x_0 is an initial state (which is treated as a parameter of the model). Equivalently, we could describe the same system in terms of the so-called innovations representation

$$y_k = F_k^*\left(y_{[0,k-1]}, u_{[0,k]}; x_0\right) + e_k, \tag{2.5}$$

where now the F_k^* define the conditional mean $\mathbb{E}[y_k \mid y_{[0,k-1]}, u_{[0,k]}]$ (assuming it exists). The F_k^* define an optimal predictor for the system (2.4), with e_k being the part of the output y_k that cannot be predicted [76].

In case the noise e_k appears additively at the output of the system, then (2.4) becomes

$$y_k = F_k\left(u_{[0,k]}; x_0\right) + e_k \tag{2.6}$$

in which case we have $F_k = F_k^*$. The model (2.6) is called an *output-error* model [77].

The general framework of the prediction error method allows the ground truth system to be in a feedback loop with an adaptive feedback element, given by

$$u_k = H_k\left(y_{[0,k-1]}, u_{[0,k-1]}, r_k\right), \tag{2.7}$$

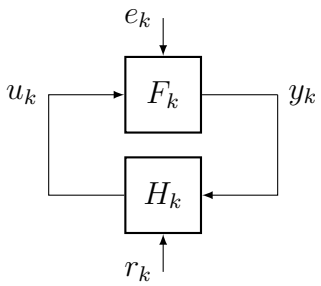


Fig. 2.11 Standard PEM system identification setup. The signals u_k and y_k are known, and used to identify the system F.

whose inputs are given by y_k and an external signal $r_k \in \mathbb{R}^{n_w}$. In this thesis, the external signal r_k is always treated as known. Figure 2.11 shows a block diagram representation of the resulting feedback system[3].

The identification problem is one of predicting y_k based on the measured quantities $y_{[0,k-1]}$ and $u_{[0,k]}$. To solve this problem, we rely on a parametric predictor model, which is a system given by

$$\hat{y}_k(\theta) = \hat{F}_k\left(y_{[0,k-1]}, u_{[0,k]}; \theta\right), \tag{2.8}$$

where $\theta \in \mathcal{D} \subseteq \mathbb{R}^{n_\theta}$ denotes a vector of parameters, with n_θ the number of parameters. When dealing with a ground truth system in the output-error form (2.6), the predictor model is given by

$$\hat{y}_k(\theta) = \hat{F}_k\left(u_{[0,k]}; \theta\right), \tag{2.9}$$

in which case it is common to call the predictor a simulation model.

In the prediction error method, optimal parameter estimates $\hat{\theta}$ for the predictor model are found based on N collected input-output data points, given by the sequences $y_{[0,N]}$ and $u_{[0,N]}$. The estimates are found based on a criterion function that needs to be minimized; many such criteria exist, so that the PEM encapsulates many different kinds of estimation methods.

A simple criterion (corresponding to the maximum likelihood method if the noise is Gaussian-distributed) that can be used to obtain parameter estimates is the mini-

[3]In (2.5) and (2.9), we allow the input to affect the output without a delay. This differs from the text in [76], where a time delay is assumed. However, as remarked by [76], this delay is not essential for the results. To make clear that algebraic loops are not allowed in the system, we included an explicit time delay in the subsystem (2.7).

mization of the mean-square prediction error

$$\varepsilon_k(\theta) = y_k - \hat{y}_k(\theta)$$

Thus by minimizing

$$V_N(\theta) = \frac{1}{N}\sum_{k=1}^{N} \|y_k - \hat{y}_k(\theta)\|^2, \tag{2.10}$$

we obtain the parameter estimates

$$\hat{\theta}_N = \arg\min_{\theta\in\mathcal{D}} V_N(\theta) \tag{2.11}$$

The asymptotic behavior of the parameter estimates $\hat{\theta}_N$ as the number of data points N tends to infinity depends on the asymptotic behavior of the function $V_N(\theta)$. The asymptotic analysis of parameter estimates is concerned with studying these behaviors. Usually, the analysis is divided into two parts: convergence and consistency. Here, we discuss a general criterion for convergence, first stated in [76].

Since the system is stochastic, $V_N(\theta)$ is a random variable. To guarantee that the identified model is independent of the specific realization of the noise entering the system, we need the prediction error $\varepsilon_k(\theta)$ to satisfy an ergodicity property: $V_N(\theta)$ must converge to its expected value as $N \to \infty$. This property is achieved by means of two fundamental assumptions: one on the ground truth system, and one on the model structure.

Assumption 1 ([76], [40])**.** The closed-loop system (2.4) and (2.7) is such that for each $k, s \in \mathbb{Z}_+$, $k \geq s$, there exist random variables $\bar{y}_{k,s}$ and $\bar{u}_{k,s}$, independent of $r_{[0,s]}$ and $e_{[0,s]}$ but not independent of $r_{[0,k]}$ and $e_{[0,k]}$, such that

$$\mathbb{E}\left[\|y_k - \bar{y}_{k,s}\|^4\right] < C\alpha^{k-s} \tag{2.12a}$$

$$\mathbb{E}\left[\|u_k - \bar{u}_{k,s}\|^4\right] < C\alpha^{k-s} \tag{2.12b}$$

for some $C > 0$ and $\alpha < 1$, with $\bar{y}_{t,t} = \bar{u}_{t,t} = 0$.

Assumption 2 ([76])**.** The mappings $\hat{F}_k$ are differentiable with respect to θ for all $\theta \in \mathcal{D}$. Let $\mathcal{D}$ be closed and bounded. Assume there exist a $C < \infty$ and $\alpha < 1$ such that

$$\left\|\hat{F}_k\left(y_{[0,k-1]}, u_{[0,k]}; \theta\right) - \hat{F}_k\left(\tilde{y}_{[0,k-1]}, \tilde{u}_{[0,k]}; \theta\right)\right\| \leq C\sum_{m=0}^{k}\alpha^{k-m}\left(\|u_m - \tilde{u}_m\| + \|y_m - \tilde{y}_m\|\right) \tag{2.13}$$

and

$$\left\| \hat{F}_k \left(0_{[0,k-1]}, 0_{[0,k]}; \theta \right) \right\| \leq C \tag{2.14}$$

for all u, $\tilde{u}$, y, $\tilde{y}$, and θ that belongs to an open neighborhood of $\mathcal{D}$. Assume also that the $(d/d\theta)\hat{F}_k$ are subject to (2.13).

When the model (2.8) or (2.9) satisfies Assumption 2, then the mapping $\theta \mapsto \{\hat{F}_k(\cdot, \cdot; \theta)\}_{k \in \mathbb{Z}_+}$ is called a *model structure* [77, Section 5.7]. Thus (2.8) and (2.9) are called model structures when viewed as a function of θ.

The main convergence result of [76] can now be stated as follows.

Lemma 1 ([76])**.** *Consider the feedback system* (2.5)-(2.7) *subject to Assumption* (1)*, and the model* (2.8) *subject to Assumption* (2)*. Consider* $V_N(\theta)$ *given by* (2.10)*. Then*

$$\sup_{\theta \in \mathcal{D}} |V_N(\theta) - \mathbb{E}\left[V_N(\theta)\right]| \to 0 \qquad \textit{w.p. } 1 \textit{ as } N \to \infty$$

If the result of Lemma 1 holds, then we have a guarantee that the criterion function $V_N(\theta)$ obtained with a specific realization of our measurements will grow independent of the specific realization of the noise (and of the input, if the input is a stochastic process) as the number of gathered data increases.

Fading memory and convergence of the cost function

Assumptions 1 and 2, which are instrumental in the convergence result of Lemma 1, are tightly connected to the fading memory properties discussed in Section 2.2.1.

To explore this idea, suppose for simplicity that the identification setup is the open-loop identification of a single-output system; in this case, $u_k = r_k$ and y_k are scalars. Then, the condition (2.12a) of Assumption 1 can be thought of as a form of stochastic approximately-fading memory property. To see this, we can define an operator F such that $y_k = (F[u, e]^\top)_k = F_k \left(u_{[0,k]}, e_{[0,k]}; x_0 \right)$. With $u_k = r_k$, it then follows that Assumption 1 can be rephrased in terms of the inequality

$$\mathbb{E}\left[|(F[r, e]^\top)_k - (FW_{k,k-s}[r, e]^\top)_k|^4 \right] < C\alpha^{k-s}$$

where $W_{k,k-s}$ is the window operator defined in (2.3). The window function discards the elements $r_{[0,s]}$ and $e_{[0,s]}$, which should not affect the variable $\bar{y}_{k,s}$ for the required independence to hold. In a purely deterministic case (without the expectation operator), the inequality above defines a strong kind of approximately-finite memory property, in

which convergence of the two outputs is exponential with respect to the length of the time-window $W_{k,k-s}$.

Note that the same relationships with the approximately-finite memory property can be established in the closed-loop case (where $u_k \neq r_k$). In that case, we have to deal with the two system operators that map r_k and e_k to y_k and u_k.

It can similarly be shown that Assumption 2 directly implies fading memory of the predictor model. This will be shown for output-error models in Chapter 4.

Fading memory and parameter optimization

We have just seen that an (exponentially fast) fading memory property ensures that the effect of noise in the cost function $V_N(\theta)$ becomes less important as the number of collected data points increases. However, even in a completely deterministic setting, a fading memory is important for tractably estimating the parameters of a nonlinear model structure.

As observed by [2, 1], when the dynamics of the model structure possesses positive Lyapunov exponents[4], solving the optimization problem (2.11) may become intractable even when no noise is present in the measurements. In particular, instabilities caused by positive Lyapunov exponents severely hinder the search for optimal parameters θ, due to the creation of a large quantity of local and possible spurious minima in the parameter landscape.

The relevance of fading memory in this context comes from the fact that exponential contraction of the model structure dynamics precludes the existence of positive Lyapunov exponents. As mentioned in Section 2.2.2, exponential contraction can be understood as a state-space version of the fading memory property. Thus, a fading memory facilitates the search for the optimal parameter estimates in (2.11).

Recent work

Despite the fact that assumptions 1 and 2 are only sufficient for the consistent estimation of nonlinear models with the PEM, these conditions have not been relaxed since the publication of [76].

For instance, in [98], a block-oriented linear-fractional transformation (LFT) model structure is used for the identification of nonlinear systems with process noise using the PEM. This model structure allows the authors to consider a more flexible setup than the closed-loop setup of Figure 2.11, at the cost of only considering the estimation of

[4]For a brief exposition to this concept, see Section 5.1.1 in Chapter 5.

static nonlinearities within the model. The authors rely on an exponential stability assumption similar to Assumption 2 to ensure convergence and consistency of parameter estimates in their method.

As remarked by [20], the PEM should be applied with caution when identifying systems using sampled limit cycle data. This follows from the fact that unforced systems with limit cycles do not satisfy Assumptions 1 and 2 (remember that, as seen in Figure 2.10, such systems do not have a fading memory on a space of uniformly bounded inputs). In [20], the authors study the application of the PEM in a context where the system F_k in Figure 2.11 is LTI, and the system H_k is a known static nonlinearity. The authors assume that when $r_k = e_k = 0$, the system has a limit cycle, and their objective is to identify the LTI system when the limit cycle is only perturbed by noise, i.e., e_k is filtered white noise, and $r_k = 0$. In their convergence analysis, the authors bypass Assumption 1 and Lemma 1 by directly assuming that all signals in the system are mean-ergodic, i.e., their time averages tend to their expectations. The question of which classes of limit-cycling systems would satisfy this assumption is not tackled, and as far as we know this is still an open question.

In [85], the authors remark that there are no generally applicable methods of system identification (or model reduction) for oscillatory systems. This further illustrates that the fading memory property (not held in general by nonlinear oscillators) is a standing assumption in system identification theory.

2.3.2 Nonlinear identification starting from linear models

There are many situations in which it is desirable to obtain a linear approximation of a nonlinear system, and this led many authors to propose frameworks under which well-defined linear approximations exist [35, 115, 83, 78].

The approaches pursued by those authors are similar in nature, but differ in the underlying assumptions on the ground truth nonlinear system that needs to be approximated. One approach, based on the PEM, dispenses entirely with assumptions on the ground truth system, assuming only that the input and output signals are quasistationary [35, 78]; a second approach, based on the Fréchet derivative, deals with the properties of the ground truth system in a case-by-case basis (in particular, Wiener systems and finite-memory systems are considered) [83]; and a third approach assumes that the output of the ground truth system (to inputs of a certain class) can be approximated arbitrarily well in mean-square sense by the output of a uniformly bounded Volterra series [115, 125].

The common thread linking these approaches is that the linear approximations are only meaningful if the nonlinear system being approximated has properties which are either similar to, or implied by, the fading memory property. For instance, for a nonlinear system to map quasistationary inputs to quasistationary outputs (an implicit assumption in [78]), there seem to be heavy constraints on its memory properties, as the few existing results on this matter have only been obtained for fading memory systems such as finite-memory and Wiener and Hammerstein systems [84]. Similarly, the existence of approximating Volterra series can be directly related to the fading memory property, as discussed in the previous section.

The Best Linear Approximation (BLA)

The linear approximation framework described in [115, 125] has enjoyed great popularity for the purposes of nonlinear system identification, as it can be efficiently used to solve the nonlinear problem using an initial linear approximation of the system. These authors work with the assumption that the system inputs belong to a certain class (denoted $\mathcal{S}_E$) of asymptotically normally distributed input signals. This class includes most notably Gaussian coloured noise and sums of sinusoids with uniformly distributed random phases – a type of signal known as random phase multisines. It is also assumed that the nonlinear ground truth system belongs to the class of systems which, when excited by signals in the class $\mathcal{S}_E$, can have their outputs approximated arbitrarily well, in mean-square sense, by the outputs of uniformly bounded Volterra series. The linear approximation, in this case, is called the Best Linear Approximation (BLA) of the system. The ground truth BLA, which can be of either parametric (a real-rational transfer function) or non-parametric (a frequency response) form, can be consistently identified by exciting the nonlinear system with an input in the class $\mathcal{S}_E$, and then using a linear model structure and a least-squares cost function to identify the system.

The success of the BLA framework for nonlinear system identification has been well established on two classes of model structure: block-oriented models and polynomial nonlinear state-space models.

Block-oriented models are those in which the system can be described as an interconnection between a finite number of LTI systems and a finite number of static nonlinearities. This type of structure includes classical models such as Wiener and Hammerstein models [142], which are SISO systems[5], but it also includes MIMO systems given by the parallel interconnection of Wiener and Hammerstein models [128].

[5]In Wiener models, the LTI block is at the input of the system, and it is cascaded with a static nonlinearity at the output; in Hammerstein models, the blocks are organized the other way around.

As reviewed in [129], identifying the BLA of a block-oriented system is a good first step in the nonlinear identification procedure. The reason is the fact that using signals in the class $\mathcal{S}_E$ yield a BLA that is a simple function of the linear blocks of the system[6].

An alternative to block-oriented models is the polynomial nonlinear state-space model structure (PNLSS) – a state-space description where the vector field is given by polynomial functions. It has been shown that extracting a linear state-space model from the nonparametric BLA of the nonlinear system (by means of linear subspace methods) can provide good initial parameter estimates for the nonlinear identification problem [100]. Using a PNLSS model structure, [26] studied the problem of identifying the forced Van der Pol oscillator model structure and an identification procedure relying on the estimation of the BLA as an initial step. Despite the fact that the estimation of a BLA for forced autonomous oscillators has no theoretical justification, for high input power the authors were able to find a stable parametric BLA that could be used to initialize the nonlinear optimization of the nonlinear model. The authors also provided a heuristic analysis for informative input design, showing that random-phase multisines have an advantage over other forcing signals since the multisines drive the ground truth system' states more uniformly around the phase-space.

2.4 Identification of neuronal systems

To discuss the identification of neuronal systems, it is useful to review the main types of neuronal models that are currently used in neuroscience. Limiting ourselves to *spiking* (as opposed to *rate-based)* single-compartment models, we can classify[7] the most popular models in the following way:

- Biophysical models: these models are descendants of the squid giant axon model of Hodgkin and Huxley [57] (see Figure 2.2). Models in this class tend to have multiple nonlinear components that can be interpreted in neurophysiological terms. In a deterministic setting, biophysical models are usually defined by a smooth *state-space* model. The two main paradigms of biophysical models are conductance-based models [63, 37, 66] and Markov models [66, 56, 19, 74].

- Generalized linear integrate-and-fire (GLIF) models: these models are descendants of the one-dimensional integrate-and-fire model introduced by Lapicque [70] (see

[6] If the ground truth system is a Wiener system, for instance, then its BLA is equal to the linear element of the system, modulo a real-valued gain [142].

[7] Our classification is based on [63, 47, 14, 5].

also [3]). In this model class, the subthreshold membrane voltage dynamics is modeled using linear time-invariant *filters*, and a discontinuous mechanism is used both for the generation of spikes and for resetting initial conditions after spiking. In a deterministic setting, spiking occurs when the voltage reaches a hard threshold; in a stochastic setting, spiking is more likely to happen as the threshold is approached. Two modern paradigms in this model class are Linear-Nonlinear Poisson models [25, 101] and Generalized Linear Models (GLMs) [114, 46].

- Simple nonlinear models: these models are descendants both of the FitzHugh-Nagumo model [38], originally developed to analyze neuronal dynamics in the phase plane, and of the Lapicque integrate-and-fire model. In this model class, a small number of simple nonlinearities are used to model the membrane voltage dynamics, and the number of state variables is kept at two or three at most. While the initiation of spikes is entirely based on equilibrium bifurcations in the nonlinear dynamics, some models in this class use a reset mechanism to implement spike downstroke (see., e.g., [46, Chapter 5], [62, 148]).

Recent research on the estimation of neuronal models from experimental data has tended to focus on the two first classes of models above. Simple models are limited in their ability to quantitatively reproduce real neuronal data, and are thus mostly used for qualitative numerical simulations and mathematical analysis.

In the next two sections, we describe biophysical and GLIF models in more detail, and review recent research on the estimation of such models from input-output data. We then finish the section by discussing the complementary roles that these modeling frameworks play in our efforts to better understand the brain.

2.4.1 Biophysical models

Biophysical models can be thought of as generalizations of the Hodgkin-Huxley model (see Figure 2.2). These models are in general given by the voltage dynamics

$$c\dot{v} = -g_L(v - \nu_L) - \sum_{j=1}^{n_c} i_j + i \tag{2.15}$$

and the ion current dynamics[8]

$$\dot{x}^{(j)} = f^{(j)}(v, x^{(j)}) \tag{2.16a}$$

$$i_j = \bar{g}_j \, p_j(x^{(j)}) \, q_j(v) \tag{2.16b}$$

for $j = 1, \ldots, n_c$, with n_c the number of ion current types (e.g., sodium, potassium, calcium, etc.).

In the equation above, the scalar term $\bar{g}_j q_j(v(t))$ models the maximum electrical current that can flow through a population of ion channels on a given area of neuronal membrane — that is, the current that would flow if all ion channels in the population were open. Thus $\bar{g}_j$, called the maximum conductance, has units of conductance per square area, while $q_j(v(t))$ has units of voltage. The term $p(x^{(j)}(t)) \in (0, 1)$ represents the proportion of open channels in the population [66], or, equivalently, the probability that a single ion channel is open at time t [74]. The state variables in the vectors $x^{(j)}$ are known as *gating variables* or *kinetic states*. Understanding the phenomenological nature of gating variables is key to understanding the problem of identifying biophysical neuronal models, and thus we briefly describe where they come from.

The model (2.15)- (2.16) defines one *compartment* of a neuronal model [68]. In *single-compartment* models, the spatial properties of a neuron are ignored, and the volume encapsulated by the neuronal membrane is collapsed to a single point. In this case, there is only one membrane voltage associated to the neuron, and it evolves according to (2.15)-(2.16). In *multi-compartment* models, the topology of a neuron is not ignored, and multiple instances of (2.15)-(2.16) are used to model the evolution of the neuronal membrane in different parts of the neuron. In this thesis, we mostly deal with single-compartment models.

Kinetic ion channel modeling

Ion channels are proteins embedded in the neuronal membrane that can fold into different conformational states. One or more of these conformations allow ions to flow through the ion channel, inward or outward with respect to the neuronal cytoplasm. In other words, the channel is conducting when the protein is folded in a specific way. Equation (2.16a), together with the mappings $p(x^{(j)})$, define an ion channel's *gating dynamics*. In the tradition initiated by Hodgkin and Huxley, to describe the gating

[8]This formulation does not take into account Calcium-activated currents. Those will be discussed in an example of Chapter 5.

dynamics of ion channels, the different conformational states of an ion channel protein are modeled as a finite number of states in a *kinetic model*.

In kinetic models, the transitions between states are described by voltage-dependent transition probabilities, also called transition rates. Transition rates are usually assumed to depend only on the voltage and on the current state of the channel, so such models fall into the category of Markov models [56, 19, 74]. In this framework, each element $x_i^{(j)}$ of the vector $x^{(j)}$ in (2.16) represents the probability that the j^{th} channel is in the i^{th} conformational state. The probabilities, which must verify $\sum_i x_i^{(j)}(0) = 1$, evolve according the the so-called Master Equation,

$$\dot{x}^{(j)} = A^{(j)}(v)x^{(j)} \tag{2.17}$$

where the elements of $A^{(j)}(v)$ contain the voltage-dependent transition rates, and where $A_{i,i}^{(j)}(v) = -\sum_{k\neq i} A_{k,i}^{(j)}(v)$ for all i, j and v.

By making simplifying assumptions on the topology of the kinetic model (i.e., on how the conformational states are interconnected) and on the transition rates (e.g., some of them might be identical), a kinetic model can be greatly reduced in terms of number of states. Hodgkin-Huxley-type models have a simplified type of kinetics where

$$\dot{x}_i^{(j)} = \alpha_i^{(j)}(v)(1 - x_i^{(j)}) - \beta_i^{(j)}(v)x_i^{(j)} \tag{2.18}$$

and $p(x^{(j)}) = \left(x_1^{(j)}\right)^a \left(x_2^{(j)}\right)^b$ or $p(x^{(j)}) = \left(x_1^{(j)}\right)^a$, with a and b positive integer exponents. The terms $\alpha_i^{(j)}(v)$ and $\beta_i^{(j)}(v)$ above are the transition rates of two-state independent kinetic subunits. In Hodgkin-Huxley-type models, a further assumption is made that current conduction in a single channel is *ohmic*, in which case

$$q_j(v) = v - \nu_j \tag{2.19}$$

with ν_j the so-called Nernst potential of the j^{th} ion channel.

In the neuroscience community, models with the simplified Hodgkin-Huxley type kinetics (2.18) are known as *conductance-based models*, while models with conformational states coupled through (2.17) are known as *Markov models*. Note that despite this nomenclature, any given conductance-based model has an equivalent Markov model [56, Chapter 18], while Markov models can still be said to be conductance-based: in both types of models, the conductance is given by $n_j p(x^{(j)})$ in (2.16b).

This kinetic modeling paradigm is phenomenological, since one can hardly *prove* that a postulated model for a given channel provides a mechanistic description of

how the protein changes its conformation. At the same time, because the behavior of proteins is ultimately explained by chemical reactions occurring at very small scales, and because chemical reactions are well-described by kinetic models, one could argue that at the limit of infinitely many conformational states, a kinetic model could come close to providing that mechanistic description. Following this point of view, kinetic models are a finite-dimensional approximation of the actual dynamics of an ion channel. The view of kinetic models as approximations of a continuum of state transitions means that, as measurement technology improves and the bandwidth of ion current signals that we are able to process increases, the complexity of kinetic model structures that is required to accurately estimate the system also needs to increase [56].

It is valid to question at this point if our initial claim that ion channel system operators have a fading memory (say, on a set of uniformly bounded inputs) holds true for conductance-based and Markov models of ion channels. The fact that it is true for conductance-based will be shown[9] in Chapter 4.

Noise in biophysical neuronal models

Neurons are biological systems, and as such they are inherently noisy. This can be observed experimentally: when the same stimulus is presented to a given neuron repeatedly, the measured voltage signal, including locations of spikes along the time axis, changes in each repetition of the experiment [32, 92, 46]. This occurs because, independently of how noise is affecting the system, the noise realization changes in every experiment, and thus the output changes as well.

Noise can be taken into account in biophysical models through the introduction of noise terms in the nonlinear state-space equations. The authors in [49] have reviewed the most common ways of adding process noise to conductance-based and Markov models. There are three common noise setups:

- Input-additive noise: in this case, a noise term $e(t)$ is added to the applied current $i(t)$. This can be used both to model unmeasured, external noise (e.g., the inputs of nearby neurons in an experiment *in vivo*), or an unmeasured lumped internal current noise.

- Conductance noise: in this case, noise represents uncertainty over the proportion of open ion channels of a certain type in the membrane. It can be modeled by

[9] In fact, conductance-based models have an exponentially contracting dynamics, as shown in Example 3 in Chapter 4. This implies that they have a fading memory on any set of uniformly bounded inputs (see Corollary 1 in Chapter 4).

modifying (2.16b) as follows:

$$i_j = \bar{g}_j \left[p_j(x^{(j)}) + e(t)\right](v - \nu_j)$$

where $e(t)$ is the noise term.

- Gating kinetics noise: in this case, a noise term is added to the dynamics of the gating variables (or transition probabilities). For instance, in a conductance-based model, we have

$$\tau_i^{(j)}(v)\,\dot{x}_i^{(j)} = -x_i^{(j)} + x_{\infty,i}^{(j)}(v) + e_i^{(j)}(t)$$

where $e_i^{(j)}(t)$ is the noise affecting the i^{th} gating variable of the j^{th} ion channel.

In addition, as illustrated in [74], in Markov models, ion channel noise can be taken into account by considering that the membrane contains a finite number of ion channels, and that each individual ion channel independently transitions between its conformational states according to the transition rates in (2.17).

In Chapter 4, we will explicitly deal with the issue of input-additive white noise in the estimation of neuronal models.

Problems in biophysical model estimation

The literature on the estimation of deterministic biophysical models is focused on solving two related problems:

1. Estimation of full neuronal models: in this case, the objective is to identify a (single-compartment or multi-compartment) neuron using the model (2.15)-(2.16) and data composed of applied current and voltage traces, $i(t)$ and $v(t)$. See, e.g., [71, 60, 33, 32, 45]. This problem is usually referred to as "automatic parameter extraction" in the neuroscience community.

2. Estimation of models for ion channels: in this case, the objective is to identify a particular ion channel using a kinetic model and data composed of macroscopic ion current and voltage traces, $i_j(t)$ and $v(t)$. See, e.g., [19, 93]. This corresponds to estimating (2.16) for a single index j. Obtaining this type of data is much more laborious than obtaining data for the first problem, since, to obtain macroscopic current data for a single channel type, it is necessary to block other channel types pharmacologically.

A strategy ubiquitously explored in the literature to solve the first problem above is to make use of ion channel models that have been previously obtained by solving the

second problem. In fact it was this very strategy that allowed Hodgkin and Huxley to postulate the conductance-based model. This strategy works well for neuronal systems because of their modularity: given a set of ion channel kinetics that have already been estimated[10], a subset of those models can be used to construct a "grey-box" model structure used to identify the full neuronal system. This results in a model structure with not many parameters: commonly, only the maximal conductances $\bar{g}_j$ are left to be optimized. In this thesis, we will mostly be concerned with the first problem above.

Methods for estimating biophysical neuronal models

There are two main challenges in the estimation of full biophysical neuronal models. First, there is the fact that neither the neuron being identified (the true system) nor the full model (2.15)-(2.16) have a fading memory on spaces of commonly used inputs. This is a simple consequence of the fact that neurons may display limit-cycles and chaotic behavior; see Figure 2.10 for a simple illustration. Second, it is difficult to apply classical estimation frameworks (such as the PEM) directly to the problem, and as a consequence analyzing the properties of parameter estimates (such as convergence and consistency) may become intractable.

A popular[11] parameter estimation method is that of [32]. In this method, the authors rely on a multi-objective optimization technique where different cost functions are constructed based on various features of the neuronal behavior such as spiking rate, spike width, spike overshoot, and spike latency. These features are obtained from data generated by applying steps of applied current to a real neuron. The resulting optimization problem is solved with a genetic algorithm. Note that this approach is radically different from the PEM approach reviewed in Section 2.3.1. By working with spiking features, which are averaged over multiple realizations of the identification experiment, the authors avoid using a voltage-based prediction error cost function.

A different estimation method, much closer in spirit to the PEM, is that of [60]. In this method, the authors make two important assumptions: first, that we can estimate the capacitive current $c\dot{v}$ from the data, and second, that the model structure contains previously identified ion channel kinetics (see previous section). They then make the important observation that it is possible to estimate the maximal conductances $\bar{g}_j$ of (2.16) by working with the voltage v as an input, and the internal current $i - c\dot{v}$ as an output. The resulting estimation problem becomes a simple least-squares problem with

[10]There exist online databases dedicated to collecting ion channel kinetic models, for instance ModelDB [90].

[11]For instance, this method was chosen to derive models for the Allen Brain Atlas project [50].

a linear-in-the-parameters model structure. As we shall see later, in control theoretic terms, this approach is tractable because it works with a model structure given by the inverse dynamics of a neuronal system — which has a fading memory.

In the method developed in [2, 1], the issue of a lack of fading memory is dealt with through the use of a stabilizing feedback term in the model structure. This strategy is used to ensure that the cost function becomes smooth with respect to the parameters. In the limit of infinite gain in the feedback term, this strategy is equivalent to that of estimating the mapping from voltage to current, as in [60]. In [143, 97], the authors estimate both the state trajectories and the unknown parameters of neuronal models based on voltage-current (input-output) traces. To do this, they formulate a nonlinear optimization problem with a regularized least-squares objective function where the constraints are given by the discrete-time dynamics. They treat each state variable, at each sample time, as a free variable of the optimization problem. This generates a very large parameter space and a set of many thousand equality constraints (the size of which grows according to the number of measurement samples). To solve such a large problem, they rely on *sequential quadratic programming* methods [9].

2.4.2 Generalized linear integrate-and-fire (GLIF) models

While in biophysical neuronal models spikes are smooth trajectories of a nonlinear system, in GLIF models spikes are instantaneous events triggered by the output of a linear system. The underlying assumption in GLIF modeling is that information in the brain is mostly conveyed by spike timings, and that, since spikes are very brief stereotypical events, we can approximate spikes by impulse functions located at *points* in time.

In a deterministic GLIF model, such impulses are triggered whenever the membrane voltage reaches a sharp firing threshold, while in a stochastic (noisy) GLIF model, the impulses are the outcomes of a point process. Because the triggering mechanism does not require a nonlinear subthreshold dynamics to work, the subthreshold voltage in a GLIF model evolves according to a linear system.

Given the relevance of GLIFs in the neuroscience literature, we review this class of models in more detail below. This is also done in the hope that the sophisticated estimation methods developed for GLIFs can bring insight into the problem of estimating biophysical neuronal models.

Spike response model (SRM)

In a deterministic setting, most GLIF models can be represented in terms of the Spike Response Model (SRM) [92, 46]. We briefly describe this model in discrete-time form. The output of a SRM is a sequence of discrete-time impulses[12]

$$s_k = \sum_{k_s \in S} \delta_{k-k_s}, \qquad k \geq 0, \tag{2.20}$$

representing a *spike train*, with the set S containing all spike times. In the SRM, the subthreshold voltage v_k is represented separately from the spike train s_k. Given a sharp threshold $\vartheta \in \mathbb{R}$, a spike time k_s is defined by

$$k_s \in S \iff v_{k_s} \geq \vartheta, \tag{2.21}$$

that is, a spike occurs when the voltage v_k crosses the threshold ϑ.

The subthreshold membrane voltage dynamics in the SRM is given by

$$\begin{aligned} v_k &= \sum_{n=0}^{k} g_{k-n}(\theta) u_n + \sum_{n=0}^{k-1} h_{k-n}(\theta) s_n + v_{\text{rest}} \\ &= \sum_{n=0}^{k} g_n(\theta) u_{k-n} + \sum_{n=1}^{k} h_n(\theta) s_{k-n} + v_{\text{rest}}, \qquad k \geq 0, \end{aligned} \tag{2.22}$$

where u_k is an external input signal ($u_k = 0, \forall\, k < 0$), s_k is the spike train defined by (2.20)-(2.21), v_{rest} is the resting potential, and h_k and g_k are impulse responses of causal ($h_k, g_k = 0, \forall\, k < 0$) and stable linear time-invariant discrete-time systems, parameterized by a vector θ. The constraint $h_0 = 0$ ensures that the closed-loop model (2.20)-(2.22) is well-posed.

At this point it is useful to introduce some notation that will also be used elsewhere in this thesis. Let the $\mathcal{Z}$-transform of h_k (defined in some region of convergence $|z| \geq \alpha > 0$) be given by

$$H(z) = \sum_{k=0}^{\infty} h_k z^{-k},$$

and let q denote the forward-shift operator, i.e., $qv_k = v_{k+1}$ and $q^{-1}v_k = v_{k-1}$. Using this notation, we can rewrite (2.22) as

$$v_k = G(q\,;\,\theta)u_k + H(q\,;\,\theta)s_k + v_{\text{rest}} \tag{2.23}$$

[12] The discrete-time impulse signal δ_k satisfies $\delta_k = 1$ if $k = 0$, and $\delta_k = 0$ if $k \neq 0$.

where $G(q\,;\,\theta)$ and $H(q\,;\,\theta)$ are to be understood as formal power series in q^{-1}.

In the SRM framework, the linear filter $G(q\,;\,\theta)$ can be used to model the dynamics of the dendritic tree of a neuron, where the integration of pre-synaptic inputs occurs [104, 92, 46]. More generally, this filter can be used to model the integration of external stimuli over a much longer neural pathway. In this case, the filter $G(q\,;\,\theta)$ can incorporate information about the *receptive field* of the neuron being modeled [25, 114, 73]. The filter $H(q\,;\,\theta)$ can be thought to model internal ionic currents that underlie the rapid decrease in voltage (*hyperpolarization*) and the refractoriness that occurs following a spike. More practically, the filter $H(q\,;\,\theta)$ provides a mechanism for setting v_{k_s+1} to a value well below the threshold ϑ when a spike occurs at $k = k_s$. It is this last feature which makes the SRM a generalization of simpler integrate-and-fire models.

Noise in GLIF models: the Generalized Linear Model (GLM)

The SRM given by (2.20)-(2.23) defines a deterministic nonlinear system. To model the stochasticity of neuronal behavior, localized noise terms (such as those discussed in the context of biophysical models in Section 2.4.1) could be added directly to the SRM equations. In that case, because the subthreshold dynamics is linear, under certain noise assumptions, it can be possible to derive an analytical expression for the likelihood of a random spike train s_k. This in turn allows for efficient maximum likelihood estimation of model parameters [106].

Stochasticity in GLIFs can alternatively be taken into account by directly defining the likelihood of a spike train given the input to the model. This idea leads to Generalized Linear Models (GLM) of neuronal spiking. Similarly to the deterministic SRM, the output of a GLM is also a train of spikes s_k, given by (2.20). However, in a GLM, the set of spike timings, S, is determined by the realization of a *stochastic point process*. In the literature, it is common to approach modeling from a continuous-time perspective, and define the GLM by means of an inhomogeneous Poisson process [46, 104, 72]. For estimation purposes, the model is discretized using a simple discretization scheme.

We will briefly describe the GLM approach in a fully discrete-time context. In this case, given a sampling period Δt, we assume that Δt is small enough so that at most one spike can be found between sampling times. A spike occurs at the k^{th} time bin if it occurs in the interval $[\,k\Delta t, (k+1)\Delta t\,)$, in which case $s_k = 1$. The probability that $s_k = 1$ is defined to be

$$\mathrm{p}(s_k = 1) = \rho_k \Delta t,$$

where

$$\rho_k = f(v_k - \vartheta) = f(G(q\,;\,\theta)u_k + H(q\,;\,\theta)s_k + v_{\text{rest}} - \vartheta) \tag{2.24}$$

is an instantaneous firing rate. The function $f(\cdot)$ and the threshold ϑ shape the probability of firing in terms of the instantaneous value of the subthreshold voltage v_k. The voltage in turn depends on the model's input and spiking history according to the linear filtering model encountered previously in the SRM.

Since (for sufficiently small Δt) the only alternative to $s_k = 1$ is $s_k = 0$ (no spike at k), a spike is a Bernoulli trial. The conditional probability mass function (PMF) of s_k, given all past spikes and given the external input, is given by

$$\mathrm{p}(s_k \,|\, s_{[0,k-1]}, u_{[0,k]}) = (\rho_k \Delta t)^{s_k}(1 - \rho_k \Delta t)^{1-s_k} \tag{2.25}$$

In the discrete-time setting above, it is important that $f(\cdot)$ be defined so that $\rho_k \Delta t \in [0, 1]$, since the latter is a probability value. In practice, $f(\cdot)$ is often defined as

$$f(v_k - \vartheta) = \tfrac{1}{\tau} e^{\beta(v_k - \vartheta)} \tag{2.26}$$

so as to model an exponential increase in the likelihood of firing as the voltage approaches (and crosses) the firing threshold ϑ [46]. However, the unboundedness of $f(\cdot)$ might lead to an ill-defined probability. To solve this issue, the fact that $\rho_k \Delta t$ is the first-order Taylor approximation of $1 - e^{-\rho_k \Delta t}$ (at zero) can be used: $1 - e^{-\rho_k \Delta t}$ lies inside $[0, 1]$, when f is given by the exponential above. The conditional PMF (2.25) can thus be redefined as

$$\mathrm{p}(s_k \,|\, s_{[0,k-1]}, u_{[0,k]}) = (1 - e^{-\rho_k \Delta t})^{s_k}(e^{-\rho_k \Delta t})^{1-s_k} \tag{2.27}$$

where now the instantaneous probability $1 - e^{-\rho_k \Delta t}$ is well-defined for any non-negative $f(\cdot)$. The joint PMF of the spike train $s_{[0,k]}$, conditioned on the external input, can be computed using the product rule of conditional probabilities. It is given by

$$\mathrm{p}(s_{[0,k]} \,|\, u_{[0,k]}) = \mathrm{p}(s_0 \,|\, u_0) \prod_{n=0}^{k} \mathrm{p}(s_n \,|\, s_{[0,n-1]}, u_{[0,n]}) \tag{2.28}$$

where the conditional probabilities on the right-hand side are given by either (2.25) or (2.27).

Estimation of GLM parameters

Given an observed spike train $s^{\circ}_{[0,k]}$ and a known input sequence $u^{\circ}_{[0,k]}$, the parameters of the GLM can be estimated by treating (2.28) as a likelihood function, and performing maximum likelihood (ML) estimation. The ML estimate is found by solving

$$\max_{\theta,\vartheta} \; \mathrm{p}(s^{\circ}_{[0,k]} \,|\, u^{\circ}_{[0,k]}), \tag{2.29}$$

where, without loss of generality, we have assumed that v_{rest} is fixed to some reasonable resting potential value.

The discrete-time ML estimation method outlined above has the advantage that only spike timings are necessary to estimate θ and ϑ. The voltage v_k is an auxiliary variable of the GLM model, and v_k must not necessarily match the actual membrane voltage of the neuron that has generated the observed spikes [105]. Intra-cellular recordings, necessary for measuring the subthreshold membrane voltage, can thus be avoided.

Many refinements of this basic maximum likelihood procedure have been proposed in the literature. For instance, intra-cellular recordings of the membrane voltage can also be used so as to include more information in the estimation procedure. As suggested by [105], given intra-cellular observations v°_k of the subthreshold voltage, in addition to the observations $s^{\circ}_{[0,k]}$ and $u^{\circ}_{[0,k]}$, the parameters θ and v_{rest} can be estimated separately from ϑ by first solving the least-squares (LS) problem

$$\max_{\theta,v_{\mathrm{rest}}} \sum_{n=0}^{k} \|G(q\,;\,\theta)u^{\circ}_n + H(q\,;\,\theta)s^{\circ}_n + v_{\mathrm{rest}} - v^{\circ}_n\|^2, \tag{2.30}$$

which is equivalent to ML estimation when v_k is corrupted by Gaussian white noise. In this case, the GLM variable v_k is not merely an auxiliary one; it should match the actual subthreshold voltage of the neuron as close as possible. The threshold ϑ is then estimated by solving (2.29) while keeping θ and v_{rest} fixed to the LS estimates found after solving (2.30). A similar multi-step estimation strategy was used in [92], where, as an increment to the basic GLM model described in the previous sections, a linear moving threshold model was used to allow the (now time-varying) firing threshold ϑ_k to depend on the spiking history $s_{[0,k]}$.

Another important refinement of the ML procedure involves approaching the estimation problem from a Bayesian perspective, and treating the parameters as random variables themselves. When the parameters are random, the joint PMF (2.28) should be understood to be conditioned on the parameters, $\mathrm{p}(s_{[0,k]} \,|\, u_{[0,k]}, \theta, \vartheta)$. Using

Bayes' rule, the *posterior* distribution of the parameters is given by

$$\mathrm{p}(\theta, \vartheta \,|\, s_{[0,k]}, u_{[0,k]}) \propto \mathrm{p}(s_{[0,k]} \,|\, u_{[0,k]}, \theta, \vartheta)\mathrm{p}(\theta, \vartheta) \tag{2.31}$$

where $\mathrm{p}(\theta, \vartheta)$ is a *prior* distribution modeling our (subjective) belief about how the parameters are distributed, before we see any data. Point estimates can be obtained by maximizing the posterior evaluated at an observed spike train and external input; this yields the maximum a posteriori (MAP) estimate. One of the main advantages of the MAP estimate over the ML estimate is that overfitting (fitting the model to noise due to overparametrization) can be mitigated thanks to the regularizing effect of the prior in the MAP cost function [101].

Finally, for the estimation of population models, the basic GLM model can be refined by coupling it with Gauss-Markov (state-space) models which are used to model correlated unobserved common input to neurons in the population [103].

Optimal input design

The methods outlined in the previous subsection are *offline* estimation methods: they assume that all the measurements have been obtained in experiments that precede parameter estimation. A feature of offline methods is that once the external input $u_{[0,k]}$ has been selected to excite the system, it is not changed during the experiment where the data (e.g., the spike train $s_{[0,k]}$) is gathered. Even when the external input is drawn from a pre-established joint probability distribution, in the end it is a single realization that is used in the data-gathering experiment. This is implicit in the notation that we have used so far in this section, where all probability mass functions are conditioned on the external input.

As an alternative to the offline procedure, the response of the system during the experiment could be used in order to continuously guide the selection of the excitation input at future times. The idea is to avoid wasting resources on an experiment that, due to a poor choice of excitation, is not informative about the system, something that may lead to poor or biased parameter estimates. This idea underlies optimal input design methods, the core principles of which we now briefly explain.

Given the measured signals $s^{\circ}_{[0,k]}, u^{\circ}_{[0,k]}$, we would like to find a principled way of choosing the next excitation sample u_{k+1} (which is treated as a free variable in the formulation below) given the measurements. One criterion that provides such a

principled choice is to maximize the decrease in *entropy*[13]

$$\mathrm{H}(\theta, \vartheta\,;\, s^{\circ}_{[0,k]}, u^{\circ}_{[0,k]}) - \mathrm{H}(\theta, \vartheta \,|\, s_{k+1}\,;\, u_{k+1}, s^{\circ}_{[0,k]}, u^{\circ}_{[0,k]}) \tag{2.32}$$

that occurs once new data is incorporated into the posterior distribution of the parameters (2.31). Since entropy is a measure of uncertainty in the distribution of a random variable, this criterion can be shown to be a sensible one [82].

The difference in entropy (2.32) can also be interpreted as the average information about (θ, ϑ) that is gained by learning about the next output s_{k+1}, and vice versa. For this reason, (2.32) is also called the *mutual information* between (θ, ϑ) and s_{k+1} [102, 113]. Since only the second term in (2.32) depends on u_{k+1}, the criterion of maximum mutual information leads to

$$u^{*}_{k+1} = \arg\min_{u_{k+1}} \mathrm{H}(\theta, \vartheta \,|\, s_{k+1}\,;\, u_{k+1}, s^{\circ}_{[0,k]}, u^{\circ}_{[0,k]})$$

To solve the optimization problem above *on-line*, i.e., in time for the system to receive the selected input sample u^{*}_{k+1}, it is necessary to update the posterior distribution (which is used in the computation of the conditional entropy) efficiently and quickly. In [72], it is shown that the posterior can be efficiently updated by performing sequential Gaussian approximations of the distributions involved in the problem; crucial to these developments is the fact that the exponential rate-shaping function (2.26) leads to a concave log-likelihood function.

2.4.3 Why estimate?

Now that we have reviewed two important modeling paradigms of neuronal behavior, it is important to discuss how estimating a neuronal model can help in answering relevant questions in neuroscience. To underscore the importance of both biophysical and GLIF models, we discuss two research areas where one type of model is more effectively used than the other.

Control of neuronal circuits

An important question in neuroscience relates to how rhythmic behaviors in animals are generated and regulated by the nervous system. It is now widely agreed that

[13] For real-valued random variables X, Y and Z, we write $\mathrm{H}(X|Y;Z) = -\mathbb{E}_{X,Y}\left[\log_2 f_{X|Y,Z}(X|Y,Z)\right]$ for the conditional entropy of X given Y evaluated at Z, where $f_{X|Y,Z}(x|y,z)$ is the conditional density of X given Y and Z.

such behaviors originate in so-called Central Pattern Generators (CPGs), circuits of synaptically coupled neurons which are capable of autonomously producing various spiking rhythms. While researchers have been able to map complete CPGs in some species of crustaceans [86], there is growing evidence of the existence of CPGs in the mammalian [53] and human [24] spinal cord.

To study the physiological mechanisms which allow CPGs to produce different types of rhythms, the *dynamic clamp* technique [117] can be used in order to introduce virtual biophysical elements in an otherwise biological system composed of CPG neurons. In the dynamic clamp experiment, voltages from all biological neurons of a given CPG are measured, and fed to biophysical models, implemented in a computer. These models are used to compute electrical currents, which are then injected back into the biological neurons. The dynamic clamp experiment is an example of a feedback control loop, where the biophysical models play the role of a digital controller. The models in the computer can be used to probe the behavior of the closed-loop system in a controlled fashion, allowing researchers to study CPGs in detail.

Using dynamic clamp, two isolated biological neurons can be turned into a CPG by connecting them with biophysical synaptic models; in addition, virtual ionic currents can be added to each of the biological neurons so as to study the role that intrinsic currents play in circuit behavior [135, 51]. Furthermore, virtual neurons can be also included in the system using full-neuron biophysical models, effectively creating hybrid biological-digital CPGs [52].

In order to implement the virtual components of a hybrid CPG in a dynamic clamp experiment, it is necessary to use models whose dynamics faithfully reproduce the dynamics of their biological counterparts. In this context, being able to estimate neuronal models is of utmost importance: estimation helps researchers to find *nominal* models to use in their studies. Nominal models can be used as a starting point in studies where circuit behavior is probed by modification of model parameters. These observations also apply to fully computational studies, where the behavior of a specific CPG is probed entirely in the computer [27, 28].

To study CPGs using dynamic clamp and neuronal models, it is not sufficient that the models reproduce spike timings accurately. It is the dynamics resulting from coupling the nonlinearities of biophysical models to the nonlinear behavior of biological neurons which allow the different rhythms of CPGs to be robustly observed and studied.

Prediction of spike times

The problem of predicting when a neuron spikes, given certain external signals, is often called the *encoding* problem [104]. In all generality, external signals may represent signals of a completely different nature than that of the electrophysiological inputs (such as synaptic or applied currents) which are causally linked to neuronal activity. External signals can represent sensory stimuli, such as light intensity and contrast (which are visual stimuli), as well as observed motor behavior, such as hand position and velocity. In this general case, external signals are treated as extrinsic covariates of neural activity [144].

The prediction of spike timings from external signals is an important problem, since the neural code (the "language" in which the brain processes information) is believed to be at least partly based on the precise times when spiking occurs — this is the essence of *temporal coding*, for which difference schemes have been postulated [46, Section 7.6.2]. A practical application of spike time prediction can be found in the emerging field of bran-machine interfaces (see [95] for a review).

Proponents of GLIF modeling have argued that, when it comes to spike time prediction using single-compartment models, the problem of neuronal encoding has been addressed most successfully with GLIFs such as the GLM and Linear-nonlinear Poisson models [47, 14]. An important advantage of GLMs is that they can be estimated efficiently without the need for intra-cellular recordings, allowing, for instance, the estimation of whole populations of neurons from extra-cellular electrode array data [114]. In addition, the filter at the input of a GLM (the term $G(q, \theta)$ in (2.24)) can be generally casted as a spatiotemporal filter, allowing for general stimuli to be taken into account. This spatiotemporal filtering, a built-in feature of GLMs, can be used to model *receptive fields*, which are those patterns of stimuli that most excite a given neuron [25]. As a result, the GLM and its variants are state-of-the-art models of choice when it comes to spike prediction from general external signals.

When external signals are more narrowly given by electrophysiological inputs, and when intra-cellular recordings are available, both GLMs [92] and biophysical models [91] have been shown to have high predictive power. Models of nonlinear integration of inputs in the dendritic tree, which could in principle be coupled both with GLMs and with biophysical models, have also been shown to predict well the evolution of the subthreshold voltage [145].

2.5 Closed-loop system identification

In this thesis, we will approach the problem of identifying the interconnection of Figure 2.1 from the standpoint of closed-loop identification. In particular, in Chapters 4 and 5, we will work towards identifying the fading memory element within the loop. This is done, however, with the understanding that it is the closed-loop behavior that we would like to predict with the identified model.

When the fading memory element in the feedback system 2.1 is a linear time-invariant system, the identification problem becomes the classical closed-loop identification problem of control theory [40, 146, 131, 136, 48].

We start this section by reviewing the linear theory, as it provides significant insight into the nonlinear problem. We then review the main features of the nonlinear problem, using a state-space realization of the feedback structure 2.1. To conclude, we show how the linear and the nonlinear closed-loop identification problems are related.

2.5.1 Linear systems

Classical system identification theory provides insights on two related problems: how to best identify a physical system (a *plant*) inserted in an existing control loop, and, conversely, how to guarantee that a physical control loop will achieve a given performance when it is designed based on an identified plant.

Identifying a system in a control loop

The problem of identifying systems operating in closed-loop was originally motivated by control engineering applications in which it is not possible (or practical) to disconnect the plant (the system to be identified) from the control loop in which it operates. In the linear setting of Figure 2.12, there are three main approaches to tackle the problem [40, 77]:

- A direct approach, in which $u(t)$ and $y(t)$, assumed to be known, are used to identify the plant.
- An indirect approach, in which the closed-loop system is identified from the known reference $r(t)$ and output $y(t)$; the plant can then be retrieved from knowledge of $K(s)$.

- A joint input-output approach, in which the closed-loop is treated as a MIMO system with inputs $r(t)$ and $e(t)$ and outputs $u(t)$ and $y(t)$. The controller does not need to be known.

The direct approach is simple to implement, and if a model structure is used that contains the ground truth system, consistent estimates can be obtained under standard persistency of excitation assumptions [40, 77].

In the nonlinear context of this thesis, the main issue with the indirect and the joint input-output approach is the requirement of an exponentially stable predictor in the sense of Assumption 2. As we have already pointed out, this is not satisfied in general by systems with the feedback structure of Figure 2.1, as they may lack a fading memory.

In chapters 4 and 5, we will make extensive use of the direct approach in order to identify neuronal systems from this closed-loop identification point-of-view. In Chapter 5, however, we will face a situation where the ground truth system does not belong to the model structure. The next section shows that the linear framework can also shed light on that issue.

Identification for closed-loop applications

In linear control theory, it is often the case that the plant is identified with the intention of using the identified model for control design [131]. Control design is thus seen as a motivation for identification.

In a situation where consistency of parameter estimates cannot be guaranteed — for instance, because the model structure is of a lower order than the ground truth system — this problem of "identification for control" poses significant challenges. Because in that situation bias is inevitable, one needs to carefully consider in which sense the identified model should be "close" to the ground truth system. Systems which are close in one particular sense will not necessarily have similar closed-loop behaviors.

To illustrate this point, figure 2.13 shows a very simple example, taken from [7], of two simple linear systems that have open-loop frequency responses which are relatively "close" in the Nyquist diagram. Knowing that none of the systems has any unstable poles, we can tell by the Nyquist diagrams that the two systems would behave very differently in case the loop were closed with negative unity feedback: the system with the red Nyquist plot becomes unstable, whereas the system with the blue Nyquist plot remains stable. Conversely, it is also easy to find LTI systems that are very different in open loop (e.g., because one is stable and the other unstable), but have a very similar closed-loop behavior.

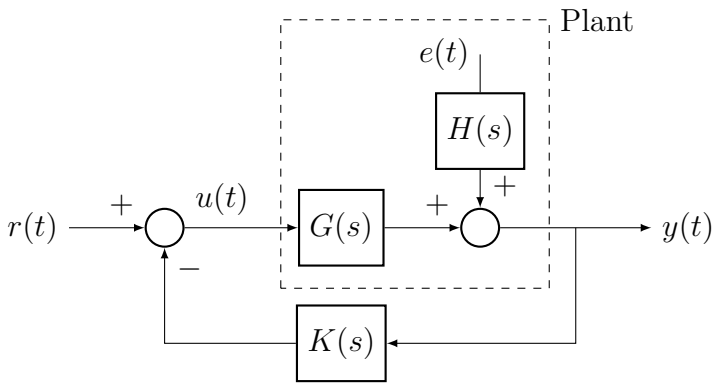


Fig. 2.12 Linear control system.

Because of those issues, several authors [147, 131, 136] have pointed out that, for predictable closed-loop operation with a controller whose design is based on an identified plant, it is crucial that the plant be accurately identified close to the crossover frequency. A crossover frequency of a transfer function $G(s)$ is any frequency ω_c such that

$$|G(j\omega_c)| = 1$$

Of particular importance is the accurate identification of the plant behavior where $G(j\omega)$ is close to the *critical point* $-1 + j0$.

In [131, Section 2.4], we can find a comprehensive analysis of the pitfalls of approximate identification for closed-loop applications. The analysis, which illustrates well the importance of identifying a plant close to the critical $-1 + j0$ point, is based on a naive identification scheme, in which an 8^{th}-order plant $G(s)$ is identified by a 5^{th}-order model $\hat{G}(s)$ by means of the frequency-domain cost function

$$\hat{G} = \arg\min_{\hat{G}} \sum_{\omega_i \in \Omega} |G(j\omega_i) - G(j\omega_i)|^2,$$

with Ω a set of logarithmically uniformly distributed frequencies.

The above cost, which puts equal weight on all bandwidths of the available data, does not necessarily yield a model $\hat{G}$ which is good for control design. Designing a controller based on an identified model which is very similar to the plant around the crossover frequency is often a condition for the real closed-loop system to maintain robust stability and performance properties of the closed-loop system with the identified plant.

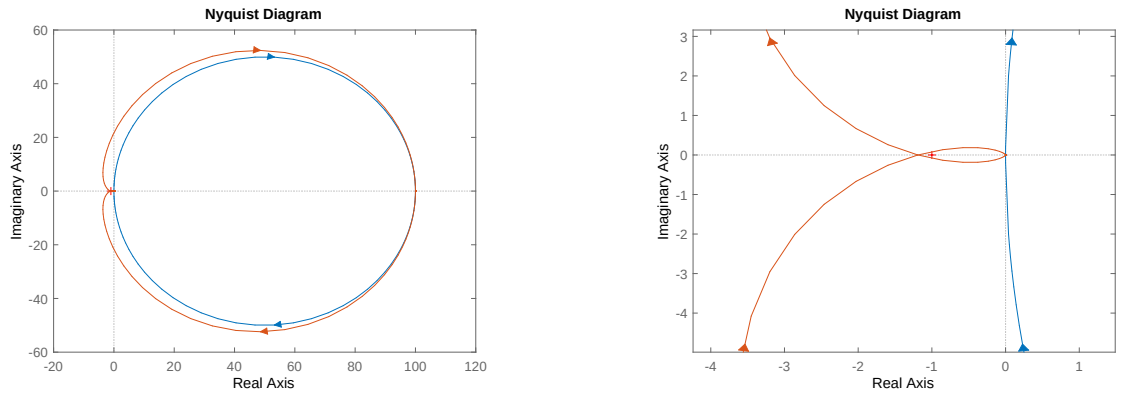


Fig. 2.13 Two linear systems with "close" open-loop responses, but very different closed-loop responses.

2.5.2 Neuronal systems

Nonlinear systems, including the majority of neuronal systems, are commonly described in terms of state-space models. In this section, we consider a state-space realization of the feedback structure of Figure 2.1. The state-space framework will allow us to discuss concepts of bifurcation theory, which are useful to understand the input-output properties of neuronal models.

The role played by the critical point $-1 + 0j$ in the classical theory of closed-loop identification, reviewed in the previous section, is tightly connected to the role played by bifurcations in neuronal state-space models. The critical point of linear control theory represents — when crossed by the Nyquist diagram — a bifurcation in a closed-loop system. The linear theory suggests that, when identifying an element within a closed-loop system using the direct method (in an "open-loop" fashion), it is important to validate the identified model well in the dynamical range that leads to a closed-loop bifurcation.

We start this section by reviewing the main types of bifurcations that occur in neuronal systems. We then show that the idea of identifying a linear system close to the critical point can be related to the idea of identifying the linearization of the neuronal model close the a bifurcation point. Finally, we relate these bifurcations to excitability classes which summarize the possible qualitative behaviors of neuronal systems. It is ultimately the excitability class of a neuronal system that must be validated by an identified model.

To simplify the discussion, in this section we deal with continuous-time systems.

State-space realization

Throughout this thesis, whenever we consider a state-space realization of the feedback structure of Figure 2.1, we assume that the system has a global normal form [61]. In this case, the system is given by the realization

$$\begin{aligned} \dot{v} &= Qx_{\mathrm{p}} - h_{\mathrm{fm}}(v, x_{\mathrm{fm}}) + b\, i \\ \dot{x}_{\mathrm{p}} &= Rv + Sx_{\mathrm{p}} \\ \dot{x}_{\mathrm{fm}} &= f_{\mathrm{fm}}(v, x_{\mathrm{fm}}) \end{aligned}$$

where R, S and Q are matrices of appropriate dimensions, b is a scalar, and f_{fm} and h_{fm} are nonlinear mappings. The passivity assumption implies that the dynamics of the linear block can be divided into the dynamics of the output $v(t)$, which is affected directly by the input $i(t)$, and the dynamics of the internal states x_{p}, which are not directly affected by the input. In other words, the passive block has a relative degree of one [18, 134].

We will work with a more concise form of the model above, given by

$$c\dot{v} = -i_{\mathrm{int}} + i \tag{2.33}$$

with

$$\dot{x} = f(v, x) \tag{2.34a}$$

$$i_{\mathrm{int}} = h(v, x) \tag{2.34b}$$

where $c = 1/b$ is a scalar, and h and f are nonlinear mappings which we assume to be smooth. We refer to (2.34) as the *internal system*, and to (2.34a) as the *internal dynamics*. The properties we discuss in this section do not depend on the particular model structure of the system (2.34).

The I-V curve

As briefly mentioned in Section 2.2.2, the operator-theoretic fading memory property is related to the state-space notion of exponential contraction. For that reason, when dealing with the state-space realization above, we will more conveniently assume that the internal dynamics (2.34a) is exponentially contracting, uniformly in the variable v.

In particular, the contraction assumption implies that the internal system (2.34) has a static input-output characteristic, also called an I-V (current-voltage) curve. This

is a mapping $i_\infty : V \to \mathbb{R}$ associating to every $v^{\text{eq}} \in V$ the unique steady-state output of (2.34) subject to the constant input $v(t) = v^{\text{eq}}$. Indeed, Proposition 2 ensures that for any constant $v^{\text{eq}} \in V$,

$$\dot{x} = f(v^{\text{eq}}, x) \tag{2.35}$$

has a unique equilibrium point, which we call x^{eq}. Thus, for every $v^{\text{eq}} \in V$, the equation

$$0 = f(v^{\text{eq}}, x^{\text{eq}})$$

has a unique solution x^{eq}. We can use these solutions to define a mapping $x_\infty : V \to X$ such that $x^{\text{eq}} = x_\infty(v^{\text{eq}})$ and $f(v^{\text{eq}}, x_\infty(v^{\text{eq}})) = 0$. The static input-output characteristic is defined by

$$i_\infty(v) = h(v, x_\infty(v)) \tag{2.36}$$

Despite its static character, the derivative di_∞/dv of the I-V curve is of particular importance to the dynamic behavior of the system, as we shall see in the next section. A closed-form expression for di_∞/dv can be obtained based on the assumption that the internal dynamics (2.34a) is exponentially contracting. In that case, it follows from Proposition 2 that the equilibrium $x_\infty(v^{\text{eq}})$ is globally exponentially stable, and so $\partial f/\partial x(v^{\text{eq}}, x_\infty(v^{\text{eq}}))$ has eigenvalues with negative real parts. This implies $\partial f/\partial x$ is invertible at $x_\infty(v^{\text{eq}})$, and thus the implicit function theorem ensures that, close to $x_\infty(v^{\text{eq}})$, we have

$$\frac{dx_\infty}{dv^{\text{eq}}}(v^{\text{eq}}) = -\left[\frac{\partial f}{\partial x}(v^{\text{eq}}, x_\infty(v^{\text{eq}}))\right]^{-1} \frac{\partial f}{\partial v}(v^{\text{eq}}, x_\infty(v^{\text{eq}})), \tag{2.37}$$

which allows finding an expression for di_∞/dv.

Equilibrium bifurcations

The set of equilibria of (2.33)-(2.34) for constant input is given by those v^{eq} that satisfy the relation

$$i_\infty(v^{\text{eq}}) = i^{\text{eq}} \tag{2.38}$$

Equilibrium point bifurcations of (2.33)-(2.34) occur when eigenvalues of the Jacobian matrix

$$J(v^{\text{eq}}) = \begin{bmatrix} -\dfrac{\partial h}{\partial v}(v^{\text{eq}}, x_\infty(v^{\text{eq}})) & -\dfrac{\partial h}{\partial x}(v^{\text{eq}}, x_\infty(v^{\text{eq}})) \\ \dfrac{\partial f}{\partial v}(v^{\text{eq}}, x_\infty(v^{\text{eq}})) & \dfrac{\partial f}{\partial x}(v^{\text{eq}}, x_\infty(v^{\text{eq}})) \end{bmatrix}, \tag{2.39}$$

cross the imaginary axis. The relation (2.38) provides the skeleton of a bifurcation diagram for the system (2.33)-(2.34), with i taken as a bifurcation parameter. The fact that i_∞ is a well-defined mapping limits the kinds of equilibrium bifurcations that can happen in the system. Two bifurcations which occur in excitable systems are shown in Figure 2.14.

Saddle-node bifurcations occur when a saddle point and a node[14] are either annihilated or created as i^{eq} is slowly changed. A necessary condition for this to happen is that a single real eigenvalue of (2.39) becomes zero, and the other eigenvalues have nonzero real parts. Since (as noted in the last section) $\partial f/\partial x$ is invertible at $(v^{\mathrm{eq}}, x_\infty(v^{\mathrm{eq}}))$, then this condition is equivalent to

$$\det\left[J(v^{\mathrm{eq}})\right] = \det\left[\frac{\partial f}{\partial x}\right] \det\left[-\frac{\partial h}{\partial v} + \frac{\partial h}{\partial x}\frac{\partial f}{\partial x}^{-1}\frac{\partial f}{\partial v}\right] = 0,$$

or (since the matrix inside the second determinant is a scalar) simply

$$-\frac{\partial h}{\partial v} + \frac{\partial h}{\partial x}\frac{\partial f}{\partial x}^{-1}\frac{\partial f}{\partial v} = 0. \tag{2.40}$$

According to (2.36) and (2.37), (2.40) is equivalent to

$$\frac{di_\infty}{dv}(v^{\mathrm{eq}}) = 0, \tag{2.41}$$

and, thus, saddle-node bifurcations may happen at stationary points of the I-V curve. For a non-degenerate bifurcation, it is further required that

$$\frac{d^2 i_\infty}{dv^2}(v^{\mathrm{eq}}) \neq 0,$$

so that those bifurcations always occur on extrema (minima or maxima) of the I-V curve. This is illustrated in Figure 2.14 (a).

Hopf bifurcations occur when an equilibrium point acquires or loses stability, absorbing or creating a limit cycle in the process. A necessary condition for this type of bifurcation to happen is that a pair of complex-conjugate eigenvalues of (2.39) cross the imaginary axis. Hopf bifurcations cannot in general be inferred from the shape of $i_\infty(v)$, as illustrated in Figure 2.14 (b).

[14] A saddle point is an equilibrium at which the eigenvalues of (2.39) are real and nonzero, and at least two of the eigenvalues have different signs. A node is an equilibrium at which the eigenvalues of (2.39) are real and nonzero, and all the eigenvalues have the same sign.

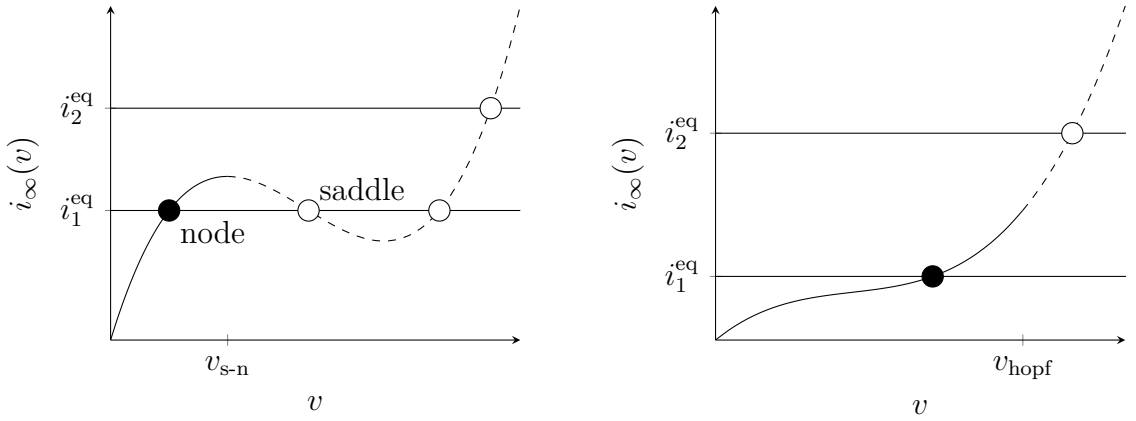


Fig. 2.14 Relationship between the I-V curve i_∞ and equilibrium bifurcations of the system. Solid lines and full circles represent stable equilibria, while dashed lines and empty circles represent unstable equilibria.

The local conductance

The function di_∞/dv is the local static conductance of the system (2.33)-(2.34), and it provides valuable information about the existence of Saddle-Node bifurcations. To obtain similar information about Hopf bifurcations, we need a dynamic version of the local conductance. This is obtained through the family of transfer functions

$$G_{\text{int}}(s; v^{\text{eq}}) = \frac{\partial h}{\partial x}\left[sI - \frac{\partial f}{\partial x}\right]^{-1}\frac{\partial f}{\partial v} + \frac{\partial h}{\partial v},$$

where the partial derivatives are evaluated at $(v^{\text{eq}}, x_\infty(v^{\text{eq}}))$. The $G_{\text{int}}(s; v^{\text{eq}})$ are the transfer functions of the internal system (2.34), when it is linearized around each of the equilibrium voltages v^{eq}:

$$\Delta I_{\text{int}}(s) = G_{\text{int}}(s; v^{\text{eq}})\Delta V(s)$$

The family of transfer functions $G_{\text{int}}(s; v^{\text{eq}})$ represent the local conductance of the system. The DC value of G_{int} is the static conductance encountered previously,

$$G_{\text{int}}(0, v^{\text{eq}}) = \frac{di_\infty}{dv}(v^{\text{eq}})$$

A family of linear systems representing the local behavior of (2.33)-(2.34) around v^{eq} is given by

$$\Delta V(s) = Z(s)\Delta I_{\text{app}}(s)$$

with

$$Z(s) = \frac{1}{cs + G_{\text{int}}(s, v^{\text{eq}})} \tag{2.42}$$

Since the roots of the denominator of the transfer function $Z(s)$ are the poles of the jacobian (2.39), it follows that a necessary condition for a Hopf bifurcation to occur at a certain certain v^{eq} is given by

$$cj\omega_h + G_{\text{int}}(j\omega_h; v^{\text{eq}}) = 0$$

for some $\omega_h \neq 0$. Thus, we can say that whenever a Saddle-Node or a Hopf bifurcation occurs at $v = v^{\text{eq}}$, we have

$$\text{Re}[G_{\text{int}}(j\omega; v^{\text{eq}})] = 0$$

for some $\omega \geq 0$. This generalizes the necessary condition (2.41), which was only valid for Saddle-Node bifurcations.

An analogy with the critical point

It is interesting to consider $G_{\text{int}}(j\omega, v)$ as a mapping from $\Omega \times V$ to $\mathbb{C}$, where $\Omega = j\mathbb{R}_+$. Then it follows that local bifurcations can only occur when $(j\omega_h, v_h)$ belongs to the set

$$\partial G = \{(j\omega, v) \in \Omega \times V \mid \text{Re}[G_{\text{int}}(j\omega, v)] = 0\}$$

The set ∂G separates regions in the the $\Omega \times V$ space associated with negative conductance, $\text{Re}[G_{\text{int}}(j\omega, v)] < 0$, and positive conductance, $\text{Re}[G_{\text{int}}(j\omega, v)] > 0$. The coexistence of these regions at different frequency-voltage ranges is a hallmark of excitable systems [30]. Knowledge of ∂G is invaluable in an identification setting because it tells us the relevant range of frequencies and voltages that should be excited in order to capture the local properties of the system.

From the discussion in the previous section, it follows that successfully identifying the linearized behavior close to the set ∂G is sufficient to accurately identify the linearized system $G_{\text{int}}(j\omega; v)/(cj\omega)$ close to the critical point $-1 + j0$.

Failing to identify this localized feature of the model may have detrimental consequences for the identification of the closed-loop model. To illustrate this point, we analyze ∂G for the two different models whose open-loop and closed-loop behaviors were previously contrasted in Figures 2.5 and 2.6. The sets ∂G of each of the models are illustrated in Figure 2.15. It can be seen that the localized regions of excitability of the two models differ significantly: in one model, the region of negative conductance

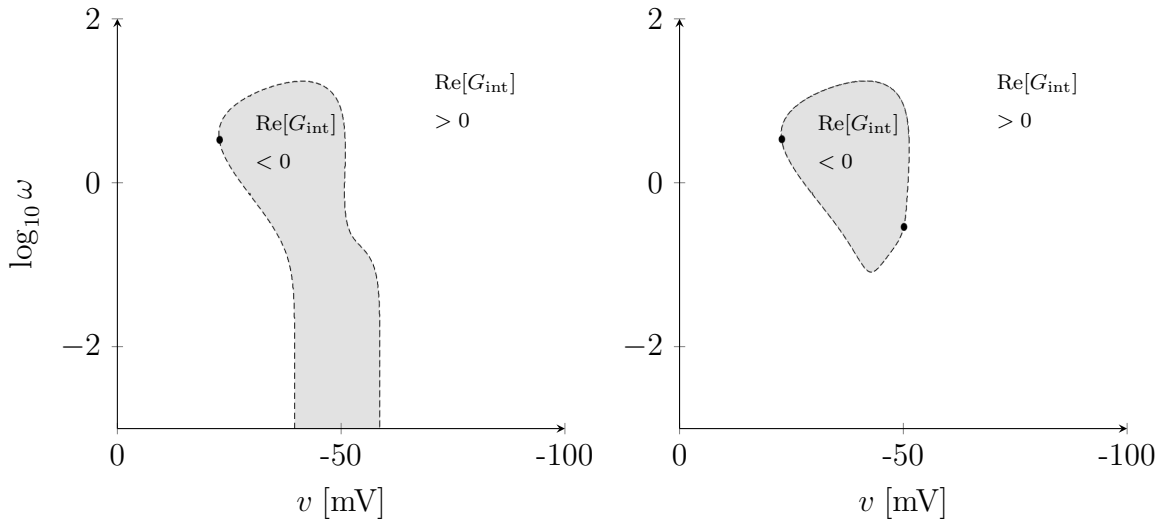


Fig. 2.15 The set ∂G (dashed) of each of the two models whose behavior was previously-contrasted in Figures 2.5 and 2.6. The set ∂G divides regions of negative conductance, $\text{Re}[G_{\text{int}}(j\omega, v)] < 0$ (grey) from regions of positive conductance, $\text{Re}[G_{\text{int}}(j\omega, v)] > 0$ (white). Black dots denote Hopf bifurcations, while Saddle-node bifurcations happen where ∂G intersects the line $\omega = 0$ (not shown).

(linked to excitability) extends into the low frequencies; in the other model, that region is bounded at higher frequencies. This shows that even though the two models are "close" in an open-loop sense (Figure 2.5), a mismatch in their linearized internal system (Figure 2.15) results in different closed-loop behaviors (Figure 2.6).

In Chapter 5, we will see various examples of models that identify the localized regions of excitability of their ground truth counterparts.

Limit cycle bifurcations

We have thus far focused on introducing local, equilibrium bifurcations in neuronal systems. Such bifurcations are relevant for our discussion since they are involved in the definition of excitability classes, introduced later in this section. Equally important to the behavior of neuronal systems are are global, limit cycle bifurcations. These bifurcations involve the creation or annihilation of limit cycles in the state-space. Among the most important limit cycle bifurcations are saddle-homoclinic bifurcations — in which a limit cycle is created from a trajectory that begins and ends in a saddle point — and fold-limit cycle bifurcations — in which a stable and an unstable limit cycle are created or annihilated analogously to saddle-node bifurcations. Discussing these bifurcations in detail is out of the scope of this work.

2.5.3 Excitability and frequency-current (F-I) curves

When a neuron that is initially at rest is subject to a an upward (or excitatory) applied current step, two behaviors may be observed: steady-state periodic spiking (Figure 2.3, top panel) and quiescence (Figure 2.3, middle panels). In case the neuron responds to the applied current step with steady-state periodic oscillations, we can associate the amplitude of the step to the oscillation frequency, $f_s > 0$. Taking into account that steady-state quiescence corresponds to a steady-state oscillation with frequency $f_s = 0$, we can thus construct a relation $f_s(i)$ between spiking frequency and applied current amplitude. This relation — which is not necessarily a mapping, as we will see — is called the frequency-current (F-I) curve.

Hodgkin and Huxley originally classified neuronal behavior using three *excitability classes* that reflected qualitative aspects of the F-I curve. Class I excitability referred to the ability of a neuron to periodically spike with an arbitrary low frequency; this behavior results in a continuous F-I curve, as shown in Figure 2.16 (left). Class II excitability referred to the inability of a neuron to fire with arbitrarily low frequencies, in which case there is a lower bound on the frequency of spiking. This behavior results in a discontinuous F-I curve, as shown in Figure 2.16 (center, black curve). When a Class II excitable neuron that is initially spiking is subject to downward (or inhibitory) applied current steps, it may be observed that the applied current value at which the neuron ceases to spike is lower than the value at which it starts to spike when subject to upward current steps. In that case, the F-I curve possesses hysteresis (Figure 2.16, center, red curve).

From the dynamical systems point of view [63, 37], Hodgkin and Huxley's classification of excitability can be reinterpreted in terms of equilibrium bifurcations in neuronal models, with the applied current taken as the bifurcation parameter. Class I excitability can be associated with saddle-node on invariant circle (SNIC) bifurcations, whereas Class II excitability can be associated with Hopf bifurcations and saddle-node bifurcations not involved in the creation of a limit cycle (sometimes called *off* invariant circle).

Until very recently, this classification of excitability[15] (whether interpreted in terms of the F-I curve shape or in terms of equilibrium bifurcations) was left unquestioned in the literature. In a series of recent papers [29, 41, 31, 42] however, some authors observed that the simple picture of excitability painted by Hodgkin and Huxley was in

[15] In addition to Class I and Class II, a third class of excitability was also originally considered by Hodgkin and Huxley to describe the behavior of neurons that could not fire periodically at all, eliciting instead a single or a few spikes in response to a stepped applied current.

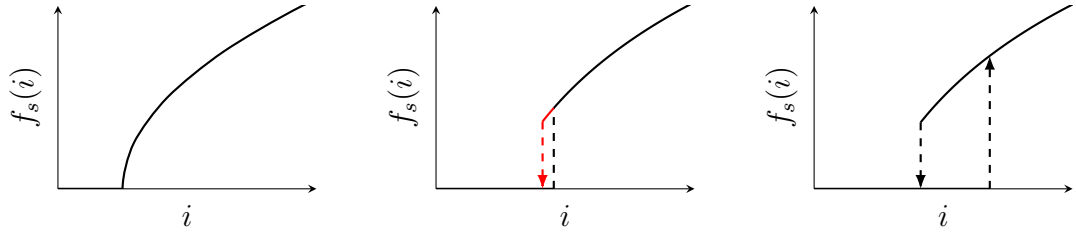


Fig. 2.16 Excitability classes and their associated F-I curves: Class I (left), Class II (center), Class II* (right) [31]. Class II neurons may have a hysteretic F-I curve, indicating rest-spike bistability; according to [42], rest-spike bistability in Class II neurons is a fragile phenomenon. This distinguishes Class II from Class II* neurons, where rest-spike bistability is robust. Class II* excitability has also been called Class IV excitability [41].

fact incomplete. For a better understanding of excitability, according to those authors, it was necessary to better understand the robustness properties of rest-spike bistability — the behavior represented by a hysteretic F-I curve.

Drion and collaborators observed that there is a fundamental difference between neuronal models with fragile and robust rest-spike bistability. According to the authors, this warranted a revision of Hodgkin and Huxley's classification scheme, with Class II excitability being broken into two distinct classes. In [31], the name Class II was retained for models with a fragile or nonexistent rest-spike bistability, while the name Class II* was used to describe models with robust hysteresis in the F-I curve. In rest-spike bistable Class II neuronal models, relatively small changes in the ratio of the timescales involved in the model may destroy the bistable behavior; in Class II* models, that behavior is robust to model uncertainties [42]. This picture of excitability is summarized in Figure 2.16.

In terms of neuronal physiology, this new classification of excitability is grounded in the idea that robust rest-spike bistability can only exist in the presence of ion currents that create a slow positive feedback mechanism in the membrane dynamics. In terms of dynamical systems theory, robust rest-spike bistability is distinguished from its fragile counterpart due to fundamentally different bifurcation mechanisms in the state-space.

Chapter 3

Feedback identification of Lure systems

In this chapter, we consider the system identification of simple feedback structures whose behaviors include nonlinear phenomena such as excitability, limit cycles and chaos. We show that output feedback is sufficient to solve the identification problem in a two-step procedure. First, the nonlinear static characteristic of the system is extracted, and second, using a feedback linearizing law, a mildly nonlinear system with an approximately-finite memory is identified. In an ideal setting, the second step boils down to the identification of a LTI system. To illustrate the method in a realistic setting, we present numerical simulations of the identification of two classical systems that fit the assumed model structure.

3.1 Introduction

As mentioned in Section 2.3.2, recent block-oriented nonlinear system identification approaches exploit the idea of estimating a best linear approximation (BLA) of the nonlinear system as a first step in the direction of solving the nonlinear identification problem. A common underlying assumption in the estimation of the BLA is that the system class has a variant of the fading memory property. The present chapter seeks to extend these methods to input-output nonlinear behaviors that can be transformed by output feedback into operators with a fading memory. More specifically, we observe that the interconnection structure in Figure 2.1 possesses that property, and is general enough to include nonlinear behaviors that are hard to identify with state-of-the art methods.

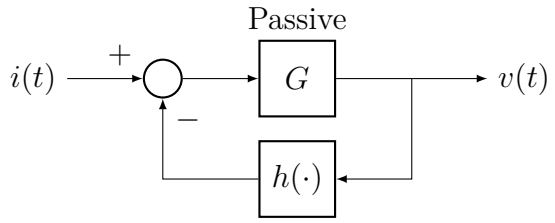


Fig. 3.1 Nonlinear Lure feedback circuit.

As a first step, we focus on the elementary situation where the fading memory component is static, and the passive element is LTI. The feedback structure then becomes the classical structure of a Lure system, shown in Figure 3.1. This simplified structure already includes famous models such as the excitable circuit of Fitzugh and Nagumo [94] and the chaotic circuit of Chua [89]. We show that the identification of such nonlinear circuits becomes straightforward if we introduce output feedback in experiment design. Not surprisingly, the static element can be identified separately from the LTI element. This allows the use of a feedback linearizing law to transform the identification problem into that of identifying a mildly nonlinear system with an approximately-finite memory [121].

Although elementary, we believe that this methodology is general and appealing for the identification of nonlinear systems that do not have the fading memory property. This methodology is also in line with the idea that smart experiment design is important to obtain good models of nonlinear systems.

The chapter is organized in the following way. In Section 3.2, we define the model class which we are interested in identifying, and give two examples of systems that fit in that class. In Section 3.3, we recall the concept of approximately-finite memory, and show that output feedback can be used to endow systems in the defined model class with that property. In Section 3.4, we develop the main contribution of the chapter, based on a two-step identification procedure for identifying systems in the model class. In Section 3.5 we present numerical simulations concerning the identification of the examples from Section 3.2 in a realistic setting. Some concluding remarks are presented in Section 3.6.

3.2 Model class

The model and input classes of interest are defined below.

Definition 3 (Model class)**.** We will work with the class of systems given by the negative feedback interconnection between a causal LTI component G and a nonlinear static map h, as in Figure 3.1. The map $h : \mathbb{R} \to \mathbb{R}$ is a continuous function such that, without loss of generality, $h(0) = 0$. In addition, there are two real constants ρ_1 and ρ_2 such that

$$\rho_1 \leq \frac{h(v_2) - h(v_1)}{v_2 - v_1} \leq \rho_2 \tag{3.1}$$

for all $v_2 \neq v_1$. The LTI component G belongs to the set of real-rational, strictly proper transfer functions $G(s) = N(s)/D(s)$ such that all poles of $G(s)$ are in $\mathrm{Re}[s] < 0$, $\mathrm{Re}[G(j\omega)] \geq 0$ for all $\omega \in \mathbb{R}$, and $G(0) > 0$.

The above implies that $G(s)$ is positive-real [67, Definition 6.4], and that $\deg D(s) - \deg N(s) = 1$. We denote $\|g\|_1 = \int_0^\infty |g(t)|dt$, where $g(t)$ is the impulse response of $G(s)$.

Definition 4 (Input class)**.** For an arbitrary $\xi > 0$, the input class $\mathcal{U} \subset \mathcal{L}_\infty(\mathbb{R}_+)$ is the set of functions u from $\mathbb{R}_+ = [0, \infty)$ to $\mathbb{R}$ such that $\sup_{t \geq 0} |u(t)| < \xi$.

3.2.1 Some examples

In this section, we provide two simple examples of circuits that belong to the model class defined above.

Example 1. The Fitzhugh-Nagumo (FHN) circuit [94], shown in Figure 3.2, was proposed as a simple model of realistic neurons, and became a paradigm of excitability. The model has the state-space representation

$$\begin{aligned} \tfrac{1}{20}\dot{v} &= -x - h(v) + i \\ \dot{x} &= -\tfrac{3}{4}x + v \end{aligned} \tag{3.2}$$

where h is given by the nonlinear characteristic[1]

$$h(v) = -v + v^3/3 \tag{3.3}$$

Note that $\rho_1 = -1$, and the nonlinear resistance is locally active (shown on the right of Figure 3.3). It can be verified that the system (3.2)-(3.3) belongs to the model

[1]Note that if $i \in \mathcal{U}$ we can always choose a bounded positively invariant state-space X for this system. Then, $h(v)$ satisfies (3.1) in X. Such a set X can be found, for instance, using the Lyapunov function $V(v, x) = v^2/2 + 10x^2$ and the standard arguments in [67, Section 4.8].

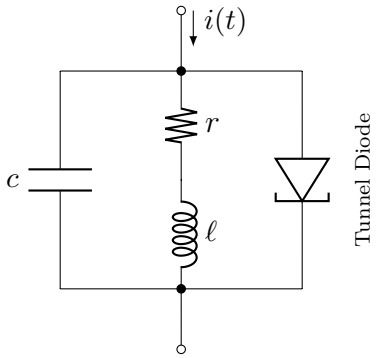


Fig. 3.2 The FHN circuit.

class of Definition 3, with

$$G(s) = \frac{20s + 15}{s^2 + 0.75s + 20}. \tag{3.4}$$

The Nyquist plot of $G(s)$ is shown on the left of Figure 3.3.

For $i = 0$, the system behaves as an autonomous relaxation oscillator, as shown in the center of Figure 3.3. For $i = -1.5$, the output $v(t)$ converges to a constant equilibrium, and the system is *excitable*: the output can display high-amplitude excursions away from equilibrium, called *spikes*, when the input i is increased momentarily past a certain excitability threshold [133].

Example 2. The Chua circuit [89] is constructed with two capacitors $c_1 > 0$ and $c_2 > 0$, an inductor $\ell > 0$, a resistor $r > 0$ and a Chua diode; it is shown in Figure 3.4. The Chua diode is a nonlinear resistive element with a piecewise-linear monotonically

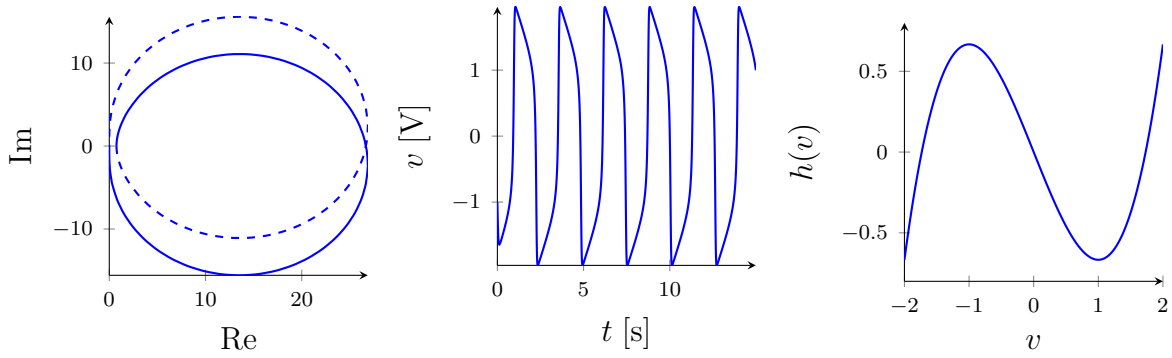


Fig. 3.3 Left: Nyquist plot of the LTI element of the FHN circuit. Center: Autonomous periodic oscillations in the FHN circuit. Right: The nonlinear element of the FHN circuit.

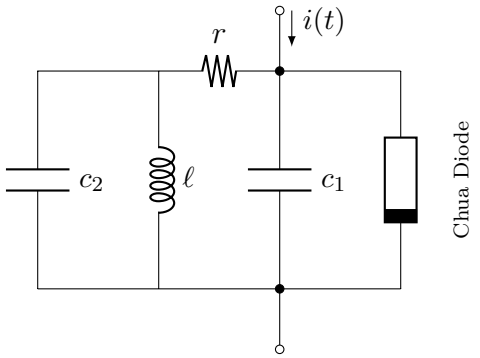


Fig. 3.4 The Chua circuit.

decreasing characteristic given by

$$h(v) = \begin{cases} -0.1(v+1)+4, & v \leq -1 \\ -4v, & -1 < v < 1 \\ -0.1(v-1)-4 & v \geq 1 \end{cases} \tag{3.5}$$

(shown on the right of Figure 3.5).

The passive element of the Chua circuit is given by

$$G(s) = \frac{\ell c_2 s^2 + \ell r s + 1}{\ell c_1 c_2 s^3 + \ell r(c_1 + c_2)s^2 + c_1 s + r},$$

and it has the Nyquist plot shown on the left of Figure 3.5.

In [89], it is shown that the autonomous Chua circuit presents chaotic behavior when the parameters are given by $c_1 = 0.1$, $c_2 = 2$, $\ell = 1/7$ and $r = 0.7$. A typical output of the system in this chaotic regime is shown in the center of Figure 3.5.

By forcing the Chua circuit with an external current, the circuit belongs to the model class of Definition 3. Note that in this case $\rho_1 = -4$ and $\rho_2 = -0.1$.

3.3 Approximately-finite memory through output feedback

In this section, we discuss how the feedback law

$$i = k(r - v) \tag{3.6}$$

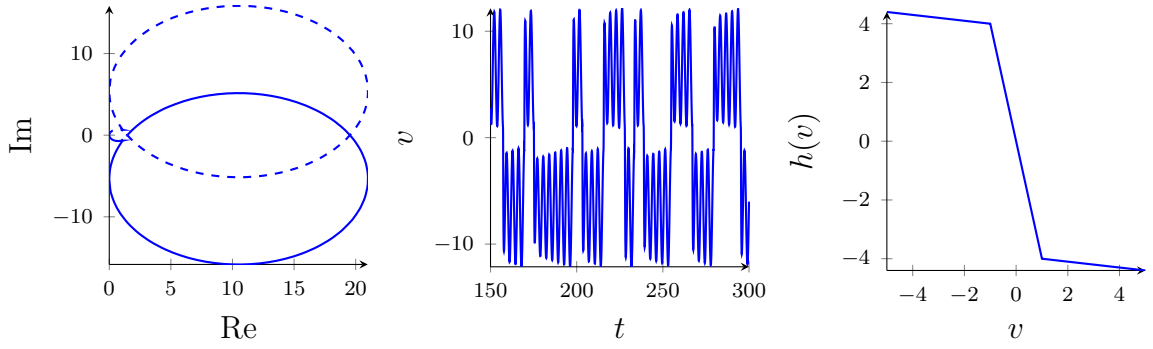


Fig. 3.5 Left: Nyquist plot of the LTI element of the Chua circuit. Center: Autonomous chaotic oscillations in the Chua circuit. Right: The nonlinear element of the Chua circuit.

is used to endow a system from the model class of Definition 3 with the approximately-finite memory property [122].

3.3.1 Approximately-finite memory

Consider the model class of Definition 3 and the input class of Definition 4. Let G denote the (convolution) operator defined by $G(s)$ and H denote the operator defined by $(Hv)(t) = h(v(t))$. It can be shown[2], based on the stability of $G(s)$ and the Lipschitz property of h, that the map $(I + GH)^{-1}$ is well defined on $\mathcal{L}_\infty(\mathbb{R}_+)$. Thus,

$$v = (I + GH)^{-1}(Gi + g_0)$$

where $g_0 \in \mathcal{L}_\infty(\mathbb{R}_+)$ is a term taking into account the exponentially decaying initial conditions of the linear system.

Let $G_{c\ell}$ denote the restriction of $(I + GH)^{-1}G$ to $\mathcal{U}$ (under zero initial conditions, this is the map from the input i to the output v). We are interested in the following property.

Definition 5 ([122])**.** Let $F : \mathcal{U} \to \mathcal{L}_\infty(\mathbb{R}_+)$ be a causal time-invariant operator. We say F has approximately-finite memory on $\mathcal{U}$, denoted $F \in \mathcal{A}(\mathcal{U})$, if for any given $\epsilon > 0$, there is a $\Delta > 0$ such that

$$|(Fu)(t) - (FW_{t,\eta}u)(t)| < \epsilon, \quad t \geq 0 \tag{3.7}$$

[2]See e.g. [122, Section 2.3], where $(I + GH)^{-1}$ is denoted by V.

for all $\eta \geq \Delta$ and all $u \in \mathcal{U}$, where $W_{t,\eta}$ is the window operator

$$(W_{t,\eta}u)(\tau) = \begin{cases} u(\tau), & t-\eta \leq \tau \leq t \\ 0, & \text{otherwise} \end{cases} \tag{3.8}$$

The inequality (3.7) shows that the recent past of the input of a system in $\mathcal{A}(\mathcal{U})$ dominates the behavior of its output. An important result linking Definition 5 to the circle criterion is [122]. In our context, we have the following statement.

Proposition 3. *Assume that one of the following two conditions are satisfied:*

(i) $0 \leq \rho_1 < \rho_2$*, all poles of* $G(s)$ *are in* $\mathrm{Re}[s] < 0$*, and* $\mathrm{Re}[G(j\omega)] \geq 0$ *for all* $\omega \in \mathbb{R}$*.*

(ii) $\rho_1 < 0 < \rho_2$*, all poles of* $G(s)$ *are in* $\mathrm{Re}[s] < 0$*, and the locus of* $G(j\omega)$ *for* $-\infty < \omega < \infty$ *is contained within the circle of radius* $(\rho_2^{-1} - \rho_1^{-1})/2$ *centered on the real axis of the complex plane at* $-(\rho_2^{-1} + \rho_1^{-1})/2 + j0$*.*

Then $G_{c\ell}$ *has approximately-finite memory on* $\mathcal{U}$*.*

Proof. Let $\mathcal{U}'$ be defined similarly to $\mathcal{U}$, but with $\xi' = \|g\|_1 \xi$. Under our assumptions, [122, Theorem 1] ensures that the map $GH(I+GH)^{-1} \in \mathcal{A}(\mathcal{U}')$ (for simplicity, we denote operators and their restrictions by the same symbols). But since $GH(I+GH)^{-1} = I - (I+GH)^{-1}$, it follows from direct application of the inequality (3.7) that $(I+GH)^{-1}$ is also in $\mathcal{A}(\mathcal{U}')$. Thus, $G_{c\ell}$ is the cascade interconnection of $(I+GH)^{-1} \in \mathcal{A}(\mathcal{U}')$ with $G \in \mathcal{A}(\mathcal{U})$. Since $Gu \in \mathcal{U}'$ for all $u \in \mathcal{U}$, $G_{c\ell}$ can be shown to be in $\mathcal{A}(\mathcal{U})$ using the cascade interconnection result[3] [121, Theorem 3]. □

3.3.2 Linear output feedback

If $h(v)$ possesses regions of negative conductance, i.e., $\rho_1 < 0$, and $G(s)$ fails to satisfy the circle condition (ii) of Proposition 3, the interconnection of Definition 3 might fail to belong to $\mathcal{A}(\mathcal{U})$ for any $\mathcal{U}$. In fact, we can argue that is the case for the two examples of Section 3.2.1. The Fitzhugh-Nagumo model, for instance, does not satisfy (3.7) for the input $i_p(t) = (\mu(t-t_1) - \mu(t-t_2))\xi/2$, with μ the Heaviside function and $t_2 > t_1 > 0$. With zero initial conditions, this input can be used to drive the state of (3.2)-(3.3) away from an unstable equilibrium at the origin and towards a stable limit cycle. As a consequence, for any constant η, (3.7) cannot hold for arbitrarily large $t > 0$ and small ϵ. A similar argument can be used for the Chua circuit, where the limit cycle is replaced with a chaotic attractor.

[3] This result requires $(I+GH)^{-1}$ to be uniformly continuous on $\mathcal{L}_\infty(\mathbb{R}_+)$, which can be shown by means of [119, Corollary 3a].

The feedback law (3.6) can be used to endow the closed-loop operator with the approximately-finite memory property. To see this, note that the closed-loop feedback system with input kr and output v can be described by the negative feedback interconnection of $G(s)$ with the static nonlinearity

$$h_k(v) = h(v) + kv \tag{3.9}$$

so that now we have

$$\rho_1 + k \leq \frac{h_k(v_1) - h_k(v_2)}{v_1 - v_2} \leq \rho_2 + k$$

for all $v_1 \neq v_2$.

Now, it is possible to make $\rho_1 + k \geq 0$ by choosing $k > 0$ large enough. Let H_k denote the operator defined by $(H_k v)(t) = h_k(v(t))$, and consider the new closed-loop operator $G_{c\ell,k} = (I + GH_k)^{-1}Gk$. Now (i) of Proposition 3 is satisfied, and we have $G_{c\ell,k} \in \mathcal{A}(\mathcal{U})$.

3.4 A feedback identification method

In this section, we show that it is possible to decouple the problem of identifying a nonlinear system belonging to the model class of Definition 3 into a nonlinear static identification stage and a dynamic mildly nonlinear identification stage. We work with the following simplifying assumption.

Assumption 3 (Simplified setup)**.** The model class is described by Definition 3. In addition, h is given by

$$h(v) = a_1 v + \sum_{j=2}^{J^*} a_j \phi_j(v) \tag{3.10}$$

where $a_j \in \mathbb{R}$ and $J^* \in \mathbb{N} \cup \{\infty\}$. We assume the ϕ_j are known linearly independent functions which are Lipschitz continuous on every bounded subset of $\mathbb{R}$. The feedback law $i = k(r - v)$, with $k + \rho_1 > 0$, is implemented with an ideal analog circuit. The signal r is known, and the signals $i_m = i + e_i$ and $v_m = v + e_v$ are observed, where e_i and e_v are mutually independent Gaussian coloured zero-mean noise terms with finite variances. Figure 3.6 with the block $K = k(r - v)$ gives a representation of this setup.

3.4.1 Static identification stage

We start by introducing the following concept.

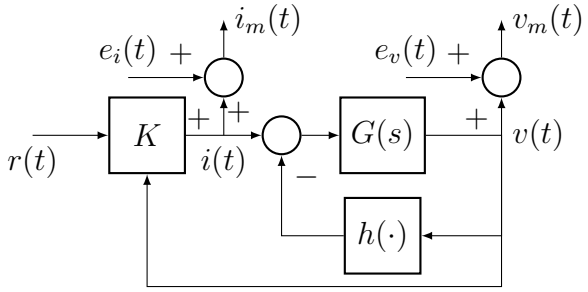


Fig. 3.6 Simplified output noise setup. $K = k(r - v)$ in the static identification stage, and $K = \kappa(r, v)$ from (3.18) the dynamic stage.

Definition 6. We define the inverse static input-output characteristic by

$$i_\infty(v) = \frac{1}{G(0)}v + h(v) \tag{3.11}$$

where $G(0) > 0$ by assumption.

The characteristic $i_\infty(v)$ gives the (unique) constant input required to establish an equilibrium at each constant v. Notice that under Assumption 3, estimating i_∞ effectively amounts to estimating the nonlinear terms of h in (3.10).

To estimate i_∞, we need to stabilize the system at different steady-states $\bar{v}$. We ensure this by means of the output feedback (3.6). The equilibrium of the system must satisfy

$$-\frac{1}{G(0)}v + kr = h(v) + kv, \tag{3.12}$$

By assumption, the right-hand side of (3.12) is monotonically increasing. Since h is continuous, it follows that (3.12) has a single solution $\bar{v}$ for every $\bar{r}$. The fact that the system settles to the unique $\bar{v}$ when subject to a constant $\bar{r}$ is guaranteed by the approximately-finite memory property [121, Theorem 2]. This can be alternatively be shown using the circle criterion [67, Theorem 7.2].

We can now discuss how to estimate i_∞. A simple procedure begins by choosing a sufficiently large $k > 0$ and a grid of M constant values for $\bar{r}$. Assume this grid is contained in the vector $\bar{V}_\mathrm{r}$. For each $m = 1, \ldots, M$, we apply the input $\bar{V}_\mathrm{r}[m]$ to the closed-loop system and wait for the system to settle to a corresponding output equilibrium $\bar{V}[m]$. This yields (as $t \to \infty$) an M-dimensional vector of true output steady-state values $\bar{V}$. In practice, the noise assumptions allow us to obtain consistent

estimates $\hat{v}$ for $\bar{v}$ and $\hat{\imath}$ for $i_\infty(\bar{v})$ by averaging the measurements,

$$\hat{v}_N = \frac{1}{N}\sum_{n=1}^{N} v_m(nt_s), \quad \hat{\imath}_N = \frac{1}{N}\sum_{n=1}^{N} i_m(nt_s) \tag{3.13}$$

where t_s is the measurement sampling period and N is the number of samples. This yields estimate vectors $\hat{V}_N$ and $\hat{I}_N$.

Considering Assumption 3, a natural estimator for i_∞ is

$$\hat{\imath}_\infty(v) = w_1 v + \sum_{j=2}^{J} w_j \phi_j(v) \tag{3.14}$$

where w_j are the estimator parameters, and $J \in \mathbb{N}$ is such that $J \leq M$. In order to estimate these parameters, we construct a matrix $\Phi_{N,J} \in \mathbb{R}^{M\times J}$ whose m^{th} row is given by

$$\left(\hat{V}_N[m], \phi_2(\hat{V}_N[m]), \ldots, \phi_J(\hat{V}_N[m])\right) \tag{3.15}$$

Assume that $\Phi_{N,J}$ has full rank. This can be accomplished by choosing a sufficiently wide and fine grid for the elements of $\bar{V}_{\text{r}}$. Then, a parameter estimate $\hat{W} = (\hat{w}_1, \ldots, \hat{w}_J)^T$ is obtained by solving

$$\min_{W} \quad \sum_{m=1}^{M} \left(\hat{I}_N[m] - \hat{\imath}_\infty(\hat{V}_N[m])\right)^2 \tag{3.16}$$

which yields

$$\hat{W}_{N,J} = (\Phi_{N,J}^T \Phi_{N,J})^{-1} \Phi_{N,J}^T \hat{I}_N \tag{3.17}$$

We thus have that, as $N \to \infty$ and $J \to J^*$, as long as $\Phi_{N,J}$ has full column rank for all J, $\hat{\imath}_\infty(v)$ converges to $i_\infty(v)$, and each $\hat{w}_j$ converges to a_j for $j = 2, 3, \ldots, J^*$ (we drop the subscripts N and J of $\hat{w}_j$ for clarity).

3.4.2 Dynamic identification stage

The main idea in the dynamic identification stage is to use the input

$$i = \kappa(v, r) \triangleq k(r - v) + \sum_{j=2}^{J} \hat{w}_j \phi_j(v) \tag{3.18}$$

so as to linearize the system by feedback.

Assumption 4. The feedback law (3.18) is implemented with an ideal analog circuit. The setup of the problem is represented by Figure 3.6, with $K = \kappa(v, r)$ given by (3.18).

From the analysis in the previous section, as $N \to \infty$ and $J \to J^*$, the identification problem becomes one of identifying a linear system with input r, output v, and output-additive Gaussian coloured noise. The ground truth model, at those limits, is given by $G_k(s) = kG_a(s)/(1 + kG_a(s))$, with $G_a(s) = G(s)/(1 + a_1G(s))$. The system $G_a(s)$ lumps together the term a_1v and the transfer function $G(s)$, which are indistinguishable from each other from the input-output perspective. The resulting linear identification problem is a well-known one for which consistency guarantees can be obtained with a variety of methods [77].

In practice, obviously, N and J will be finite, and the nonlinearity will not be perfectly canceled by feedback. In that case, identifying the closed-loop system from r to v amounts to identifying a mildly nonlinear system that has an approximately-finite memory and is subject to output-additive noise. The Best Linear Approximation (BLA) framework [125] ensures in this setting that, by using linear identification methods which are based on minimizing a squared sum of output residuals, we can obtain (asymptotically) an optimal approximation of the nonlinear system. Optimality, in this case, is defined with respect to the assumed input class [129]. Furthermore, due to the fact that operators with approximately-finite memory map periodic inputs to asymptotically periodic outputs [121, Theorem 9], by choosing periodic exciting signals, we can mitigate noise effects in the output by averaging the signal v_m over different periods.

Given a best linear estimate $\hat{G}_k(s)$, to recover the estimate of the original nonlinear system with input i and output v, we first compute

$$\hat{G}_a(s) = \frac{1}{k}\frac{\hat{G}_k(s)}{1 - \hat{G}_k(s)}, \tag{3.19}$$

which is necessary to account for the $k(r - v)$ term in (3.18). The identified nonlinear system is then given by interconnecting, in negative feedback, the transfer function $\hat{G}_a(s)$ and the nonlinearity $\hat{h}(v) = \sum_{j=2}^{J} \hat{w}_j\phi_j(v)$.

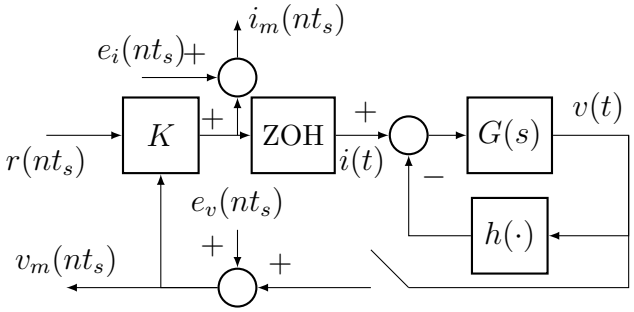


Fig. 3.7 Realistic identification setup. $K = k(r - v)$ in the static static identification stage, and $K = \kappa(r, v)$ from (3.18) in the dynamic stage. The block ZOH is a standard zero-order hold.

3.5 Simulations with a realistic setup

In a more realistic identification setting, the user-defined feedback loop around the physical system is implemented in discrete-time, and output measurement noise is fed back into the system dynamics.

In this section, using numerical simulations, we naively apply the procedure described in Section 3.4 to identify the two systems from section 3.2.1, assuming the realistic setup of Figure 3.7. We assume that e_v and e_i are given by white Gaussian noise with the same variance, denoted by σ. With this, we aim to provide a proof of concept that the method still performs well in a realistic scenario.

We briefly describe the simulation procedure. Given a vector $\bar{V}_\mathrm{r}$, each of the M experiments of the static identification stage was simulated by numerically integrating the dynamics of the scheme shown in Figure 3.7, with K given by $k(r - v)$, with the input $r(t) = \bar{V}_\mathrm{r}[m]$, $t \geq 0$, and with zero initial conditions[4]. Numerical integration was carried out for 100 seconds, which was sufficient to see (3.13) converge.

To generate data for the dynamic identification stage, we performed R simulations[4] corresponding to R realizations of two periods of the random-phase multisine inputs given by

$$r(nt_s) = \sum_{\ell=-N_f}^{N_f} u_\ell \sin(\tfrac{2\pi}{N}\ell n + \theta_\ell), \qquad n = 0, 1, 2, \ldots$$

where the θ_ℓ are random variables uniformly distributed over $[0, 2\pi[$, $N = T/t_s$ is the number of samples per signal period T, and $N_f = f_{\max}T < N/2$ is the harmonic number corresponding to the largest frequency in the signal, $f_{\max}$. The coefficients u_ℓ

[4]The simulations were performed in Matlab's Simulink™ using the numerical integration routine `ode15s` with a maximum step set to 10^{-4} seconds and relative/absolute tolerances set to 10^{-6}.

Table 3.1 Parameters used in the generation of data.

	t_s	k	$f_{\max}$	R	T	σ	SNR
FHN	10^{-3} s	1.5	100 Hz	5	500 s	0.01	40 dB
Chua	10^{-3} s	5	100 Hz	5	500 s	0.01	40 dB

are chosen such that $u_0 = 0$ and $u_\ell = \bar{u}$, with $\bar{u}$ a constant used to set the input RMS level. Simulations were carried out by numerically integrating the dynamics of the scheme shown in Figure 3.7, with zero initial conditions, and with K given by $\kappa(r, v)$ in (3.18).

Using the generated data, a continuous-time transfer function $\hat{G}_k(s)$ was estimated using the off-the-shelf Matlab System Identification Toolbox[5] routine `tfest`[6]. The number of poles and zeros of the identified transfer were constrained to be the same as those of the ground truth ones. The identified linear model is recovered as in (3.19).

The results to be discussed next were obtained with data generated using the parameters in Table 3.1. The signal-to-noise ratio (SNR) value refers to ratio of the average power of the output of the noiseless system in the dynamic identification stage, and the noise variance σ^2.

3.5.1 Fitzhugh-Nagumo circuit

Using the basis functions $\phi_j(v) = v^j$, $j = 2, 3$, Figure 3.8 shows that assuming a realistic setting results in a small error $(i_\infty - \hat{i}_\infty)(v)$. The error remains roughly the same when the noise variance is increased by a factor of 10.

Figure 3.9 shows validation of the identified model in closed-loop. For validation purposes, the mean of the input $i(t)$ was set to -1.5, which puts the FHN system in the excitable regime, and results in a characteristic spiking behavior. It can be seen that the error is kept low for most of the time, except at moments when the model "misses" a spike. These misses occur due to the ultrasensitivity of excitable systems with respect to their inputs.

3.5.2 Chua's circuit

To capture the nonlinear components of a piecewise-linear nonlinearity such as (3.5), we use the basis functions $\phi_2(v) = \max\{0, v-1\}$ and $\phi_3(v) = \max\{0, -(v+1)\}$. Figure

[5]Toolbox version 9.9, Matlab version R2018b.

[6]The function `tfest` was used with standard settings. The routine initializes parameters through the Instrument Variable (IV) method, and updates the parameters by minimizing a weighted prediction error norm using a nonlinear least-squares search method.

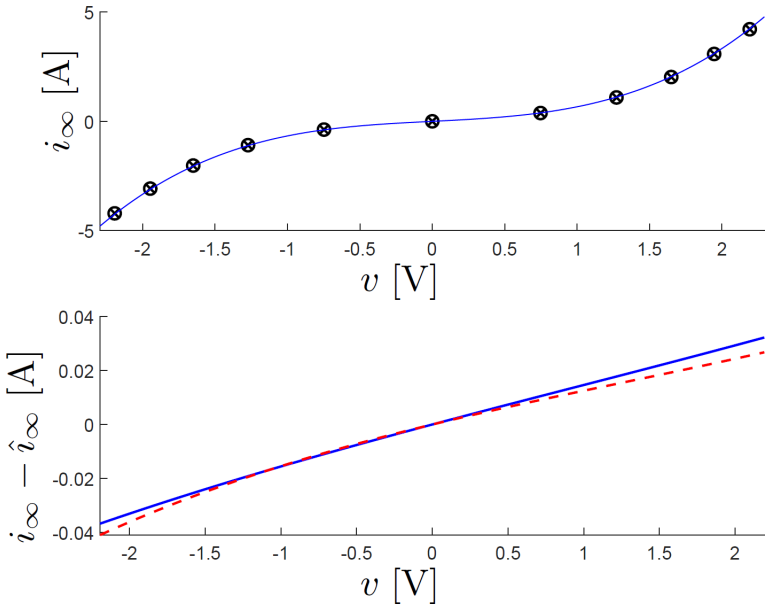


Fig. 3.8 Estimation of i_∞ for the FHN circuit. Top: ground truth i_∞ (line), estimates $(\hat{v}, \hat{\imath})$ with $\sigma = 0.01$ (crosses) and $\sigma = 0.1$ (circles). Bottom: error $i_\infty - \hat{i}_\infty$ with $\sigma = 0.01$ (solid) and $\sigma = 0.1$ (dashed).

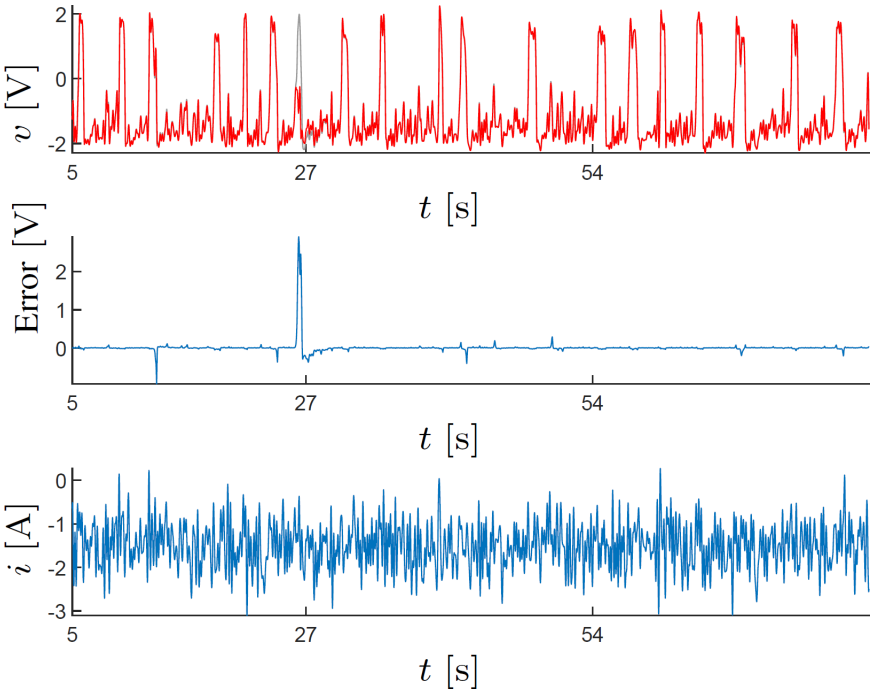


Fig. 3.9 Validation of the identified FHN circuit. Top: Ground truth model output $v(t)$ (gray) and identified model output $\hat{v}(t)$ (red). Middle: Output error $v(t) - \hat{v}(t)$. NRMSE ≈ 0.84 for the interval shown. The NRMSE increases to about 0.97 when only data from $t \geq 30$ is taken into account: most of the error comes from the "missed spike" around 27 s.

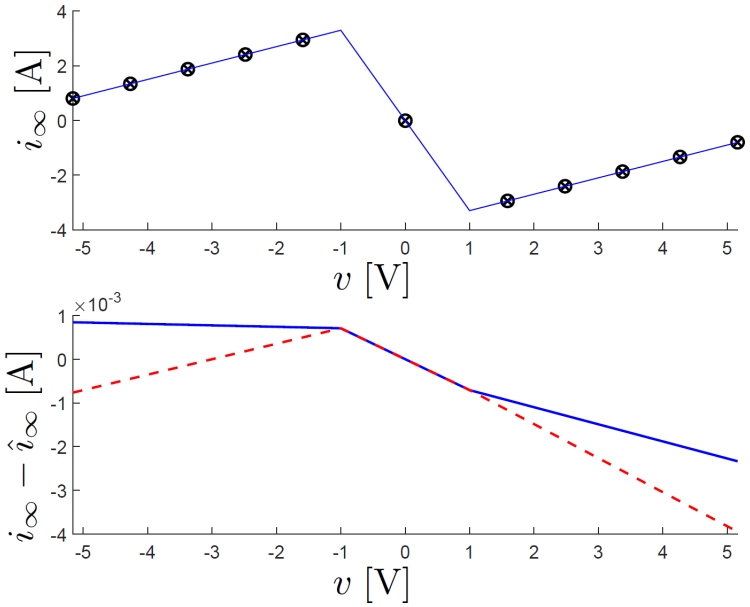


Fig. 3.10 Estimation of i_∞ for the Chua circuit. Top: ground truth i_∞ (line), estimates $(\hat{v}, \hat{\imath})$ with $\sigma = 0.01$ (crosses) and $\sigma = 0.1$ (circles). Bottom: error $i_\infty - \hat{i}_\infty$ with $\sigma = 0.01$ (solid) and $\sigma = 0.1$ (dashed).

3.10 shows the resulting nonlinearity estimation error. Again, a tenfold increase in measurement noise does not severely affect the error.

To compare the complete identified model with the ground truth model, we first realize the linear dynamics of each system (lumped with the linear component of h) in the modal canonical state-space form. Starting from a nonzero initial condition, the resulting trajectories are shown in Figure 3.11. It can be seen that the "double-scroll" attractors are qualitatively very similar.

Notice that, in this example, for illustration purposes, we have chosen basis functions which can be used to exactly describe the true nonlinearity. Had we used instead $\phi_2(v\,;\theta) = \max\{0, v - \theta\}$, with $\theta \neq 1$, for instance, the error in identifying the nonlinearity would have been higher. In practice, nonlinearly parameterized basis functions such as $\phi_2(v\,;\theta)$ can be used in the static nonlinearity estimation problem, at the cost of increased complexity in the optimization problem that needs to be solved.

3.5.3 Discussion

It can be argued that the choice of the feedback gain k is key to the success of the identification procedure developed in Section 3.4 when it is applied to the more realistic case dealt with in this section. In principle, k does not need to exceed $|\rho_1|$ by a very large margin, and indeed we chose it to be only slightly larger than $|\rho_1|$

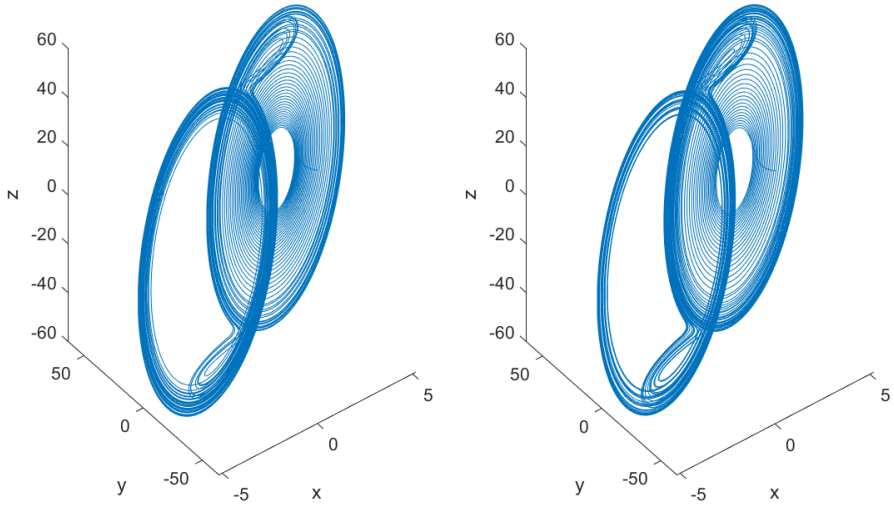


Fig. 3.11 Attractors of the ground truth Chua circuit (left) and of the identified Chua circuit (right). The trajectories in the states (x, y, z) are obtained with a modal canonical state-space realization of the original $G_a(s) = G(s)/(1 - 4G(s))$ and of the estimated $\hat{G}_a(s)$.

in both simulations above. Choosing a suitable k in this case can be viewed as part of experiment design. While our choices were good enough to avoid issues with the measurement noise that is fed back into the system, it is clear that difficulties might arise for systems with a large $|\rho_1|$. If that is the case, and if it is possible, analog feedback should be used.

3.6 Conclusion

In this chapter, we have observed that feedback can simplify the identification of a nonlinear system. We have illustrated this idea with the elementary situation where the original system is the feedback interconnection of a passive LTI system and a static nonlinearity. In this case, the use of output feedback as part of experiment design provides a straighforward solution to the problem. This procedure is sufficient to identify nonlinear behaviors such as excitability (Fitzugh-Nagumo) or chaos (Chua).

In the next chapters, we aim to generalize the method to neuronal conductance-based models [57], in which case the fading memory element is dynamic rather than static.

Chapter 4

Feedback identification of conductance-based models

In this chapter, we apply the classical prediction error method (PEM) of Section 2.3.1 to the estimation of nonlinear models of neuronal systems subject to input-additive noise. While the nonlinear system exhibits excitability, bifurcations, and limit-cycle oscillations, we prove consistency of the parameter estimation procedure under output feedback. Hence, this chapter provides a rigorous framework for the application of conventional nonlinear system identification methods to stochastic neuronal systems. The main result exploits the elementary property that conductance-based models of neurons have an exponentially contracting inverse dynamics. This property is implied by the voltage-clamp experiment, which has been the fundamental modeling experiment of neurons ever since the pioneering work of Hodgkin and Huxley. We also provide an application of the method concerning model structure selection based on the identification of linear approximations of the system's inverse dynamics.

4.1 Introduction

The estimation of models for biophysical neuronal systems is a topic that has attracted considerable interest in the scientific community over the past decades [97, 45, 60, 71, 93]. However, the asymptotic properties of published estimation methods are rarely discussed. This is understandable for models that exhibit highly nonlinear dynamics including excitable behaviors and limit cycle oscillations.

The goal of this chapter is to show that rigorous convergence results can be established in the most classical framework of the prediction error method (PEM) [76, 77]. In nonlinear system identification, the convergence and consistency analysis

of the PEM is made tractable by two well-known assumptions: (i) the true system includes output-additive noise but no input-additive noise [129, 80], and (ii) the signals are generated by a system with some form of input-output stability — for instance, a fading memory [13], input-output exponential stability [76, 98], or mean-square convergence of the output to that of a Volterra series [100, 129].

Neuronal systems fail to satisfy both of these two assumptions. First, neuronal systems are primarily subject to *input*-additive noise. This noise models the stochastic fluctuations of currents traversing the neuronal membrane. For a review of the modeling of noise in neuronal systems, see [49, 46]. Furthermore, the non-equilibrium nature of neuronal behaviors excludes any reasonable exponentialy stability or fading memory assumption.

Previous works have studied the application of the PEM under these unfavorable conditions. Under an assumption of exponential stability, but allowing for process noise, [98] developed a framework to identify block-oriented (also called structured) nonlinear systems with known LTI elements. In [20], the authors justify with dynamical systems theory the application of the PEM to identify the linear element of a Lure-type system with a limit cycle; the authors have to make an ergodicity assumption in order to prove consistency of the parameters estimates. As an alternative to the PEM, [85] developed a method based on transverse contraction analysis to identify oscillatory systems under the assumption that all states of the model are available; no noise considerations are made.

The main observation underlying the present chapter is that while the fundamental assumptions of the PEM are not verified for conductance-based neuronal models, they hold for their inverse. In other words, conductance-based models satisfy the required assumptions of PEM under high-gain output feedback. This means that neuronal systems can be identified with classical techniques by relying on the direct approach of closed-loop system identification [40]. Using contraction theory [81], we rigorously justify the use of the direct approach to consistently estimate neuronal model parameters.

We show that the closed-loop approach to the neuronal system identification problem is fully consistent with the classical voltage-clamp experiment of Hodgkin and Huxley [57]. Voltage-clamp has remained to date the key experimental methodology to derive a state-space model of a neuron. We show that there is flexibility in designing a contracting output feedback law beyond the high-gain implementation of voltage-clamp. As in previous work dealing with Lure systems [15], we advocate that feedback design

is an integral element of neuronal system identification, which makes this an attractive application of closed-loop system identification theory.

The chapter is organized as follows: in Section 4.2, we review a number of classical tools of nonlinear system identification and analysis. In Section 4.3, we introduce the general class of conductance-based models and show that they have a contracting inverse. In Section 4.4, we detail the identification of the inverse dynamics of neuronal systems with the PEM and discuss the plausibility of the required assumptions. In Section 4.6, we show how the approach can be used to inform us about which model structure to use.

4.2 Preliminaries

In this chapter, we will use the convergence result of Section 2.3.1, which deals with the identification of nonlinear systems under closed-loop with the prediction error method. We will mostly deal with output error model structures. In this section, we introduce the definition of discrete-time contracting dynamics [81], which will be used extensively in the remainder of the chapter. As briefly discussed in Section 2.2.2, this can be thought of as an stronger version of the fading memory property for state-space systems.

Notation

In this chapter, we use the notation described in Section 1.1. In addition, we will use the following notation: For arbitrary $\beta > 0$, the class of n_u-valued sequences $u : \mathbb{Z}_+ \to \mathbb{R}^{n_u}$ such that $\sup_{k\in\mathbb{Z}_+} \|u_k\|_\infty < \beta$ is denoted by $\mathcal{U}_\beta^{n_u}$ (if $n_u = 1$, we drop the superscript). We will also use the notation $\mathcal{U}_\beta^{n_u}$ to denote the set of n_u-valued functions $u : \mathbb{R}_+ \to \mathbb{R}^{n_u}$ such that $\sup_{t\in\mathbb{R}_+} \|u(t)\|_\infty < \beta$. We rely on context to make clear when $\mathcal{U}_\beta^{n_u}$ refers to a class of sequences or functions.

4.2.1 Contracting discrete-time dynamics

Neuronal systems are most commonly represented by state-space models, and so we will rely on the state-space formalism of contraction theory [81] to analyze the identification problem. We present both the discrete-time and continuous-time definitions in sequence, as they are both relevant to us.

First, consider the discrete-time system

$$x_{k+1} = f(x_k, u_k) \tag{4.1a}$$

$$y_k = h(x_k, u_k) \tag{4.1b}$$

where f and h are continuously differentiable functions, $u : \mathbb{Z}_+ \to \mathbb{R}^{n_u}$ is the input signal, $y : \mathbb{Z}_+ \to \mathbb{R}^{n_y}$ is the output signal, and $x : \mathbb{Z}_+ \to \mathbb{R}^n_x$ is the state vector. We denote by $x_k = \phi_{k,s}(u, x_s)$ the solution of (4.1a) that starts at time s and is evaluated at time $k \geq s$, when (4.1a) is subject to the input sequence $u = u_{[0,\infty]}$ and initial condition x_s.

Definition 7 ([81])**.** The discrete-time dynamics (4.1a) is said to be exponentially contracting in a positively invariant set $X \subseteq \mathbb{R}^{n_x}$, uniformly on $U \subseteq \mathbb{R}^{n_u}$, if there exist a symmetric matrix sequence $P_k(x_k) \geq \epsilon I > 0$ and a constant $\alpha < 1$ such that

$$\frac{\partial f}{\partial x}^\top P_{k+1} \frac{\partial f}{\partial x} \leq \alpha^2 P_k \tag{4.2}$$

for all $k \in \mathbb{Z}_+$, $x_k \in X$, and $u_k \in U$.

Note that when $P = I$, the above definition implies that $f(\cdot, u)$ is a contraction mapping. The definition of exponential contraction generalizes this fact by allowing for a positive-definite and state-and-time-dependent contraction metric P_k. It turns out that this generalization is useful for the analysis of interconnected systems.

For brevity, when we say that a dynamics is exponentially contracting in a set, we implicitly require that the set be positively invariant.

The result below will be instrumental in connecting the contraction property to the PEM conditions of the previous section. For simplicity, we work with a constant contraction metric.

Lemma 2. *Consider the discrete-time system* (4.1)*. Let*

$$y_k = F_k(u_{[0,k]}; x_0) = h(\phi_{k,0}(u, x_0), u_k) \tag{4.3}$$

For some $\beta > 0$, assume (4.1a) *is exponentially contracting in a convex, closed and bounded set X, uniformly on $U = [-\beta, \beta]^{n_u}$, with a constant metric $P > 0$. Then there*

are $C_1, C_2 > 0$ *such that*

$$\begin{aligned} &\left\| F_k\left(u_{[0,k]}; x_0\right) - F_k\left(\tilde{u}_{[0,k]}; \tilde{x}_0\right) \right\| \\ &\leq C_1 \sum_{m=0}^{k} \alpha^{k-m} \|u_m - \tilde{u}_m\| + C_2\, \alpha^k \|x_0 - \tilde{x}_0\| \end{aligned} \tag{4.4}$$

for all $k \geq 0$, $u, \tilde{u} \in \mathcal{U}_\beta^{n_u}$, *and* $x_0, \tilde{x}_0 \in X$.

Proof. See the Appendix 4.A.1. □

The property (4.4) has been called *exponential stability* in the context of nonlinear system identification by [76, 98]. This property is in fact a stronger version of the *fading memory* property, introduced in Section 2.2.1. To show how the concepts of contracting dynamics and fading memory are related, we have the following result.

Corollary 1. *Consider the discrete-time exponentially contracting system given by* (4.1)*, where the input* u *belongs to* $\mathcal{U}_\beta$ *and the state-space* $X \subset \mathbb{R}^{n_x}$ *is convex, closed and bounded. Define the operator* F_{x_0} *by* $(F_{x_0}u)_k = F_k(u_{[0,k]}; x_0)$*, as in* (4.3)*. Then* F_{x_0} *has a fading memory on* $\mathcal{U}_\beta$.

Proof. This follows immediately from Lemma 2: choosing $w_k = \alpha^{\frac{k}{2}}$ ensures that whenever

$$\max_{m \in \{0,1,\ldots,k\}} |u_m - \tilde{u}_m|\, \alpha^{\frac{k-m}{2}} < \delta$$

we have (after a change of variables in the sum)

$$|(F_{x_0}u)_k - (F_{x_0}\tilde{u})_k| \leq C_1 \delta \sum_{m=0}^{k} \alpha^{\frac{k}{2}} < \frac{C_1 \delta}{1 - \sqrt{\alpha}}$$

so that given $\epsilon > 0$, we can choose $\delta = \epsilon(1 - \sqrt{\alpha})/C_1$ to satisfy (2.1). □

Remark 1. For a given system, it can be easier to check the contraction property, which is based on a matrix inequality, than the fading memory property, which is based on the continuity-like condition (2.1). Furthermore, there exists a well-established theory for checking the contraction property of interconnected systems [81, 137] and, in some cases, it is possible to efficiently search for contraction metrics using LMI solvers [8]. Thus, Corollary 1 provides a useful sufficient condition to check the fading memory property of a given system.

4.2.2 Contracting continuous-time dynamics

Consider the continuous-time nonlinear system

$$\dot{x}(t) = f(x(t), u(t)) \tag{4.5}$$

where f is a continuously differentiable function, $u: \mathbb{R}_+ \to \mathbb{R}^{n_u}$ is an input signal and $x : \mathbb{R}_+ \to \mathbb{R}^n_x$ is the state vector.

Definition 8. The continuous-time dynamics (4.5) is said to be exponentially contracting in a positively invariant set $X \subseteq \mathbb{R}^{n_x}$, uniformly on $U \in \mathbb{R}^{n_u}$, if there exists a continuously differentiable symmetric matrix $P(x(t), t) \geq \epsilon I > 0$ and a constant $\lambda > 0$ such that

$$\frac{\partial f}{\partial x}^\top P + P\frac{\partial f}{\partial x} + \dot{P} \leq -2\lambda P \tag{4.6}$$

for all $t \in \mathbb{R}_+$, $x(t) \in X$, and $u(t) \in U$.

Alternatively, by writing $P = \Theta^\top \Theta$, (4.6) can be written as $\frac{1}{2}\left(F + F^\top\right) \leq -\lambda I$, with

$$F = \left(\dot{\Theta} + \Theta\frac{\partial f}{\partial x}\right)\Theta^{-1},$$

4.3 Conductance-based models under feedback

4.3.1 Conductance-based models

Conductance-based models are biophysical neuronal models that admit the circuit representation shown in Figure 4.1. While the framework of the present chapter holds for multiple-input-multiple-output models, we focus on the single-input single-output

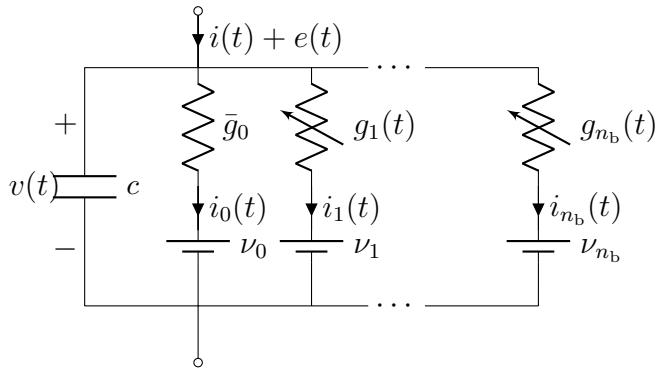


Fig. 4.1 Schematic representation of a neuronal system.

case. These models were briefly mentioned in Section 2.4.1. Here, we describe them in more detail.

In a conductance-based model, the neuronal membrane is modeled by an ideal capacitor of capacitance $c > 0$. The voltage across the membrane, which is the output of the model, is given by $v(t)$. The neuron possesses $n_\mathrm{b} \in \mathbb{N}$ different types of ion channels embedded in its membrane. The ion channels allow ionic currents to flow across the membrane according to Kirchhoff's law,

$$c\,\dot{v}(t) = -\sum_{j=0}^{n_\mathrm{b}} i_j(t) + i(t) + e(t) \tag{4.7}$$

where each current $i_j(t)$, $j = 1, \ldots, n_\mathrm{b}$, models an ionic current. The ionic currents not explicitly included in the model are lumped in a leak current $i_0(t)$. In addition, the membrane voltage is affected by an external input current $i(t)$ and an additive noise current $e(t)$. The noise current aggregates the effects of ion channel fluctuations [49] and background neuronal activity [46, Chapter 8].

All currents in a conductance-based model obey Ohm's law. The leak current

$$i_0(t) = \bar{g}_0(v(t) - \nu_0) \tag{4.8}$$

is characterized by a constant conductance $\bar{g}_0 > 0$ and a constant reversal potential $\nu_0 \in \mathbb{R}$. In contrast, the conductances of the ionic currents are voltage-dependent. This dependence is the key source of nonlinearity of conductance-based models. Owing to the original proposal of Hodgkin and Huxley, each ionic current has a nonlinear state-space model of the form

$$\dot{m}_j = \frac{1}{\tau_{m,j}(v)}\left(-m_j + m_{\infty,j}(v)\right), \tag{4.9a}$$

$$\dot{h}_j = \frac{1}{\tau_{h,j}(v)}\left(-h_j + h_{\infty,j}(v)\right), \tag{4.9b}$$

$$i_j(t) = \bar{g}_j m_j^{\alpha_j}(t) h_j^{\beta_j}(t)(v(t) - \nu_j) \tag{4.9c}$$

with $j = 1, \ldots, n_\mathrm{b}$. The constants $\bar{g}_j > 0$ are called the maximal conductances, and $\nu_j \in \mathbb{R}$ are called reversal potentials. The variables m_j and h_j are called gating variables, and take values in the closed interval $[0, 1]$. Their dynamics are defined by the continuously differentiable functions

$$\tau_{m,j}, \tau_{h,j} : \mathbb{R} \to [\tau_\mathrm{min}, \tau_\mathrm{max}] \subset \mathbb{R}_+$$

and

$$m_{\infty,j}, h_{\infty,j} : \mathbb{R} \to [0,1]$$

where $\tau_{\min} > 0$. The gating variables modulate the current conductance with a voltage-dependent first-order lag dynamics. The exponents α_j and β_j belong to $\mathbb{Z}_+$, and whenever $\alpha_{j^*} = 0$ or $\beta_{j^*} = 0$, we ignore (4.9a) or (4.9b) for $j = j^*$, respectively. These exponents, along with the gating variable dynamics (4.9a)-(4.9b), constitute the *kinetic model* of the j^{th} ion channel [66, 56].

A compact representation of the entire model (4.7)-(4.9) has the state-space structure

$$c\dot{v} = -h(v,x) + i + e \tag{4.10a}$$

$$\dot{x} = A(v)x + b(v) \tag{4.10b}$$

where the vector $x \in X = [0,1]_x^n$ collects all the gating variables m_j and h_j for which $\alpha_j > 0$ and $\beta_j > 0$, respectively, and

$$h(v,x) = \bar{g}_0(v - \nu_0) + \sum_{j=1}^{n_{\text{b}}} \bar{g}_j \, m_j^{\alpha_j} \, h_j^{\beta_j} (v - \nu_j) \tag{4.11}$$

denotes the total membrane internal current. Note that the matrix $A(v)$ is diagonal, and $b(v)$ is a vector-valued function of v. The model (4.10) is in the standard normal form of nonlinear minimum phase systems [18]: Kirchoff's equation (4.10a) is the input-output equation of a relative degree one model, and the gating dynamics (4.10b) are the system's internal dynamics.

Example 3. The Hodgkin-Huxley (HH) model [57] is the prototypical conductance-based model. It is given by (4.7), with $c = 1$ μF/cm^2, and it has two ionic currents ($n_{\text{b}} = 2$): a sodium current i_{Na}, and a potassium current i_{K}. It also includes a leak current i_{L}. The currents are given by

$$\begin{aligned} i_0 &= i_{\text{L}} = 0.3\,(v + 54.4) \\ i_1 &= i_{\text{Na}} = 120\, m_{\text{Na}}^3 h_{\text{Na}} (v - 55) \\ i_2 &= i_{\text{K}} = 36\, m_{\text{K}}^4 (v + 77) \end{aligned} \tag{4.12}$$

The three internal variables are the sodium activation $m_1 = m_{\text{Na}}$, sodium inactivation $h_1 = h_{\text{Na}}$, and potassium activation $m_2 = m_{\text{K}}$ (there is no potassium inactivation in the model, i.e., $\beta_2 = 0$). The vector x collecting these variables is given by $x = (m_1, h_1, m_2)^\top = (m_{\text{Na}}, h_{\text{Na}}, m_{\text{K}})^\top$. The different voltage dependent time-constants

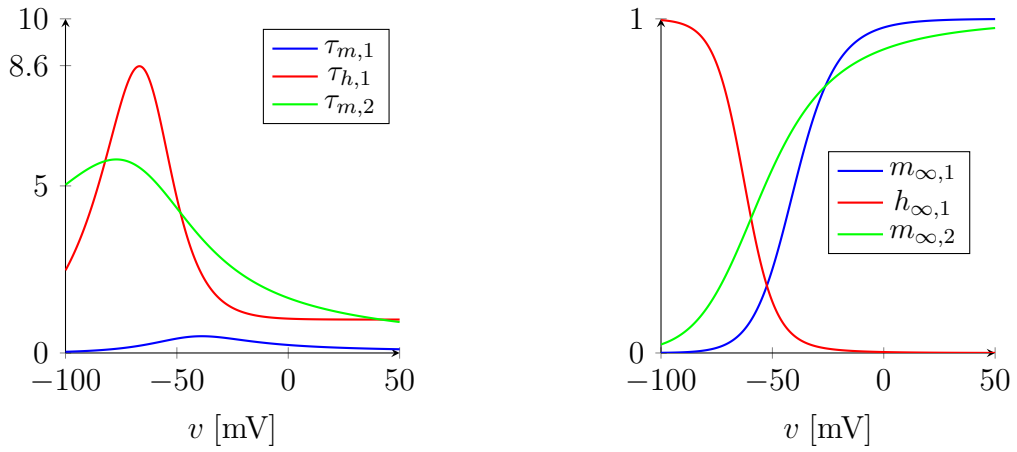


Fig. 4.2 Left: Time constant functions $\tau_{m,j}$ and $\tau_{h,j}$ of the Hodgkin-Huxley model. Right: Nonlinear activation functions $m_{\infty,j}$ and $h_{\infty,j}$ of the Hodgkin-Huxley model.

$\tau(v)$ and activation functions $m_\infty(v)$ and $h_\infty(v)$ are illustrated in Figure 4.2 and are detailed in Appendix 4.B.

Each gating variable remains in the interval $[0, 1]$, and, in the absence of external inputs, the voltage remains in the interval $[\nu_2, \nu_1] = [\nu_{\mathrm{K}}, \nu_{\mathrm{Na}}] = [-77, 55]$. This is illustrated in Figure 4.3, where a *spiking* limit cycle oscillation occurs in response to a small constant input.

The internal dynamics (4.10b) of the Hodgkin-Huxley model in Example 3 are exponentially contracting in $[0, 1]^3$, uniformly on $V = \mathbb{R}$ (Definition 8). This is verified with the constant metric $P = pI$, for any $p > 0$, and any λ such that $0 < \lambda < 1/\tau_{\max}$: in that case, we have

$$-2\,p\,\mathrm{diag}\left(\tfrac{1}{\tau_{m,1}(v)}, \tfrac{1}{\tau_{h,1}(v)}, \tfrac{1}{\tau_{m,2}(v)}\right) \leq -\tfrac{2}{\tau_{\max}} pI$$

and we could pick, for instance, $\lambda < 1/8.6$ (see Figure 4.2, left). This is in fact a general property of conductance-based models:

Proposition 4. *The conductance-based model* (4.10) *has a global relative degree of one, and is globally minimum phase (in the sense of [18]). The internal dynamics are exponentially contracting in $x \in [0, 1]^n_x$, uniformly on $V = \mathbb{R}$: there is a $P_x > 0$ and a $\lambda_x > 0$ such that*

$$P_x A(v(t)) + A(v(t))^\top P_x \leq -2\lambda_x P_x \tag{4.13}$$

for all $v(t) \in \mathbb{R}$.

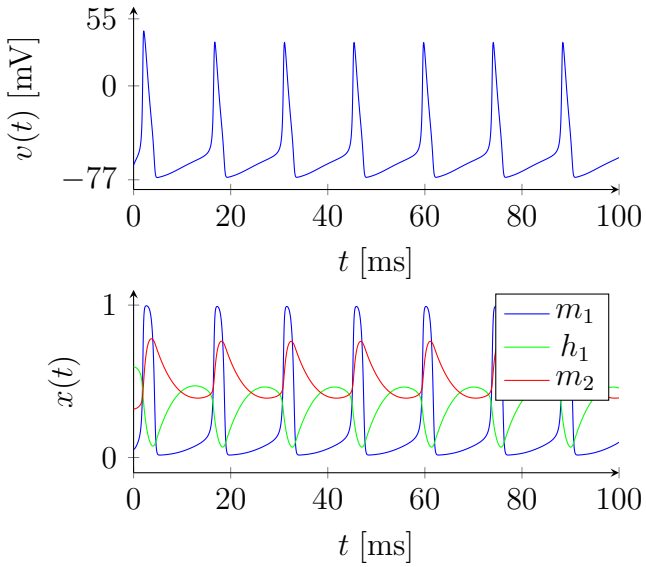


Fig. 4.3 Simulated state trajectories of Hodgkin-Huxley model (Example 3) for $i(t) = 10\ \mu\text{A/cm}^2$, $e(t) = 0\ \mu\text{A/cm}^2$.

4.3.2 Output feedback contraction

A direct consequence of Proposition 4 is that a conductance-based model has a stable inverse. More precisely, using a static output feedback law, the closed-loop dynamics can be made exponentially contracting:

Proposition 5. *Consider a conductance-based model* (4.10) *subject to the output feedback law*

$$i(t) = \gamma(r(t) - v(t)) \tag{4.14}$$

where $\gamma > 0$ *is a constant gain and* $r(t)$ *is a reference signal. Assume* $|r(t)|, |e(t)| < \beta$ *for all* $t \geq 0$. *Let*

$$V_{\beta,\gamma} = \left[\min_j\{\nu_j, -\tfrac{\beta(\gamma+1)}{\gamma}\}, \max_j\{\nu_j, \tfrac{\beta(\gamma+1)}{\gamma}\}\right] \tag{4.15}$$

Then $V_{\beta,\gamma} \times [0,1]^{n_x}$ *is a positively invariant set for the closed-loop dynamics* (4.10), (4.14), *and there exists a* $\gamma > 0$ *such that the dynamics is exponentially contracting in* $V_{\beta,\gamma} \times [0,1]^{n_x}$, *uniformly on* $[-\beta, \beta]^2$.

Proof. First, to see that $V_{\beta,\gamma} \times [0,1]^n_x$ is a positively invariant set for any $\beta, \gamma > 0$, note that $x \in [0,1]^n_x$ implies $\bar{g}_j m_j^{\alpha_j} h_j^{\beta_j} \geq 0$ in (4.9c). Then, it suffices to check that $\dot{v} \geq 0$ for $v = \min_j\{\nu_j, -\beta(\gamma+1)/\gamma\}$, and $\dot{v} \leq 0$ for $v = \max_j\{\nu_j, \beta(\gamma+1)/\gamma\}$.

We now follow an argument similar to [149, Section 2.2]. The Jacobian (with respect to the states) of the closed-loop dynamics (4.10), (4.14) is given by

$$J = \begin{bmatrix} -\frac{1}{c}(\frac{\partial h}{\partial v} + \gamma) & -\frac{1}{c}\frac{\partial h}{\partial x} \\ \frac{\partial A}{\partial v}x + \frac{\partial b}{\partial v} & A(v) \end{bmatrix}$$

(we omit dependencies on x and v for clarity). By Proposition 4, the internal dynamics (4.10b) has a contraction metric $P_x = \Theta_x^\top \Theta_x > 0$ associated with the rate $\lambda_x > 0$. We will use the matrix

$$\Theta = \begin{bmatrix} c & 0 \\ 0 & \Theta_x \end{bmatrix}$$

to define a contraction metric $P = \Theta^\top \Theta$ for the closed-loop system. Define $F = \Theta J \Theta^{-1}$ (this is the generalized Jacobian of the closed-loop system). Then

$$F = \begin{bmatrix} F_{11} & -\frac{\partial h}{\partial x}\,\Theta_x^{-1} \\ \frac{1}{c}\Theta_x \ \left(\frac{\partial A}{\partial v}x + \frac{\partial b}{\partial v}\right) & F_{22} \end{bmatrix}$$

with

$$F_{11} = -\tfrac{1}{c}\left(\tfrac{\partial h}{\partial v} + \gamma\right) \tag{4.16}$$

and

$$F_{22} = \Theta_x A(v) \Theta_x^{-1} \tag{4.17}$$

We will use $F \prec 0$ to denote $\frac{1}{2}(F + F^\top) \leq -\epsilon I$ for all $t \geq 0$, $(v, x) \in V_{\beta,\gamma} \times [0,1]^{n_x}$, and some $\epsilon > 0$. By Definition 8, to demonstrate contraction of the closed-loop system, we have to show that $F \prec 0$. To do that, we will require that $F_{11} \prec 0$ and $F_{22} \prec 0$. Contraction of the internal dynamics (see (4.13)) automatically implies

$$\tfrac{1}{2}\left(F_{22} + F_{22}^\top\right) \leq -\lambda_x I \tag{4.18}$$

and thus $F_{22} \prec 0$. Furthermore,

$$\frac{\partial h}{\partial v}(v, x) = g_0 + \sum_{j=1}^{n_\mathrm{b}} \bar{g}_j m_j^{\alpha_j} h_j^{\beta_j} \geq g_0 > 0 \tag{4.19}$$

and thus $F_{11} \prec 0$ as well. Since $F_{11} \prec 0$, by Lemma 4 in Appendix 4.D, we have $F \prec 0$ if and only if

$$\tfrac{1}{2}(F_{22} + F_{22}^\top) < Q^\top F_{11}^{-1} Q \tag{4.20}$$

where Q is the row vector given by

$$Q = \frac{1}{2}\left(-\tfrac{\partial h}{\partial x}\,\Theta_x^{-1} + \tfrac{1}{c}\left(\tfrac{\partial A}{\partial v}x + \tfrac{\partial b}{\partial v}\right)^\top \Theta_x^\top\right) \tag{4.21}$$

From (4.16), (4.19), (4.18) and (4.20), $F \prec 0$ if and only if

$$\gamma I > c\left[-\tfrac{1}{2}(F_{22} + F_{22}^\top)\right]^{-1} Q^\top Q - \tfrac{\partial h}{\partial v} \tag{4.22}$$

By (4.18) and (4.19), a sufficient condition for (4.22) to hold is

$$\gamma > \frac{c}{\lambda_x}\sigma_{\max}[Q]^2 \tag{4.23}$$

where $\sigma_{\max}[Q]$ is the largest singular value of Q.

Since the continuous functions $\partial h/\partial x$, $\partial A/\partial v$ and $\partial b/\partial v$ in (4.21) are bounded on $V_{\beta,\gamma} \times [0,1]^{n_x}$ for any $\beta, \gamma > 0$, it follows that $\sigma_{\max}[Q]$ is also bounded on that set. Since $V_{\beta,\gamma_2} \subseteq V_{\beta,\gamma_1}$ for $\gamma_2 \geq \gamma_1$, the maximum of $\sigma_{\max}[Q]$ on $V_{\beta,\gamma} \times [0,1]^{n_x}$ is non-increasing in γ. Thus, a sufficiently large γ ensures that (4.23) is satisfied. □

The expressions (4.22) and (4.23) can be used to estimate a bound on the gain that is necessary to stabilize the conductance-based model. Depending on the choice of the contraction metric P_x, this bound can of course be conservative, as illustrated by the following example.

Example 4. For the Hodgkin-Huxley model (Example 3), choose $P_x = I$ and $\lambda_x = 1/8.6 < 1/\tau_{\max}$. Consider the set $[-77, 55] \subseteq V_{\beta,\gamma}$. Computing the right-hand side of (4.22) for $v = -77$, $m_1 = h_1 = m_2 = 1$, and $\Theta_x = I$, leads to the bound 5.1×10^8 mS/cm^2 on the gain necessary to ensure exponential contraction of the closed-loop system. Alternatively, consider the contraction metric $P_x = 10^6 \times \mathrm{diag}(0.21, 3.80, 3.16)$. In this case, a random search over the set $[-77, 55] \times [0,1]^3$ gives a less conservative bound of (4.22) of 2.7×10^3 mS/cm^2.

Non-ohmic ion currents

The Ohmic expression of the ionic current (4.9c) is not necessary for the output feedback contraction property of the model. Instead of the Ohmic ionic current (4.9c), we could have used the more general formulation

$$i_j = \bar{g}_j m_j^{\alpha_j} h_j^{\beta_j} p_j(v)$$

with

$$p_j(v) = \sum_{\ell=0}^{d_j} \eta_\ell \, v^\ell$$

where each d_j is an arbitrarily large polynomial degree, and $\eta_\ell \in \mathbb{R}$. In most non-Ohmic ionic current models, $p_j(v)$ is a monotonically increasing function (see, e.g., [66, Chapter 3]), and the reversal potential $\nu_j \in \mathbb{R}$ is the value where $p(\nu_j) = 0$. In this case,

$$\operatorname{sign}(p_j(v)) = \operatorname{sign}(v - \nu_j) \tag{4.24}$$

where by convention sign $0 = 0$. Thus $V_{\beta,\gamma} \times [0,1]^{n_x}$, with $V_{\beta,\gamma}$ given by (4.15), remains a positively invariant set under the more general ionic current model above. Since the vast majority of ionic current models is Ohmic, we keep the formulation (4.9c), noting that all the results in this chapter can be easily adapted to allow for the non-Ohmic case above.

4.3.3 The voltage-clamp experiment

The output contraction property of conductance-based models is a consequence of the very experimental protocol that has been used to identify neuronal systems in the past: the voltage-clamp experiment, pioneered by Hodgkin and Huxley. The voltage-clamp experiment is nothing but a high-gain output feedback experimental protocol employed to stabilize the neuron and to determine its inverse dynamics through step response experiments. The principle of that experiment is illustrated in Figure 4.4. In the limit of high-gain feedback, the current drawn from the amplifier to clamp the voltage to the reference $r(t)$ is by definition the output of the internal dynamics driven by the voltage $v(t) = r(t)$. Electrophysiologists rely on the stability of that inverse system to model the internal dynamics through a series of step responses. In that sense, the contraction property of conductance-based models is an experimental property of neurons rather than the property of a specific mathematical model of the ionic currents.

Models of specific ion channel types have been accumulated over time by electrophysiologists. Today, online databases such as ModelDB [90] contain large libraries of ion channels models. The structure of those models is often used in parametric identification of new types of neurons (see, e.g., [32, 60]). The identified parameters include the maximal conductances $\bar{g}_j$ and the Nersnt potentials ν_j. The purpose of the next sections will be to show that, the classical PEM provides consistent estimates for these parameters.

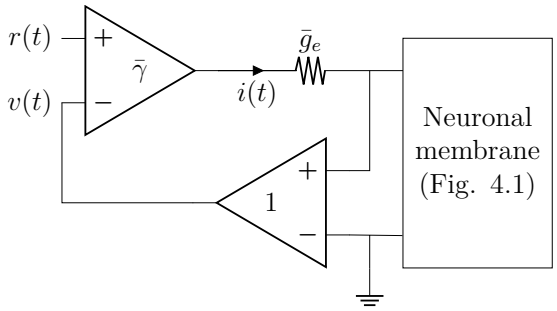


Fig. 4.4 The voltage-clamp experiment: electrodes are used to inject the current $i(t)$ and measure the voltage $v(t)$ of the neuronal membrane. The amplifiers are ideal differential amplifiers, and $\bar{g}_e$ models the electrode conductance. When $\bar{\gamma} \gg 1$, this implements the feedback law (4.14) with $\gamma = \bar{\gamma}\bar{g}_e$.

4.3.4 Discretized conductance-based models

The data obtained for identification purposes is in the form of a discrete sequence of samples. In the next sections, we will consider simple discretizations of (4.10) resulting from the forward-Euler method:

$$v_{k+1} = v_k + \tfrac{t_s}{c}\left(-h(v_k, x_k) + i_k + e_k\right) \tag{4.25a}$$

$$x_{k+1} = x_k + t_s\left(A(v_k)x_k + b(v_k)\right), \tag{4.25b}$$

where $t_s > 0$ is the sampling period.

In general, the forward-Euler method does not yield an exact discretization of a nonlinear system. In fact, such an exact discretization might not exist, as discussed in [126]. Still, some discretization methods can be more accurate than others. For instance, the authors in [152] have proposed an advanced discretization scheme for nonlinear systems that improves the order of the local truncation error with respect to simpler schemes. However, for systems with a relative degree of one, such as conductance-based models, the scheme proposed by [152] and the forward-Euler method are equivalent. This reassures us to use forward-Euler discretization in order to study the properties of the discrete-time identification problem. For that purpose, we will rely on the next result.

Lemma 3. *Assume that the continuous-time dynamics* (4.5) *is exponentially contracting in a closed and bounded set X, uniformly on $[-\beta, \beta]^{n_u}$ (with $\beta > 0$), with a constant metric P. Let*

$$x_{k+1} = x_k + t_s f(x_k, u_k) \tag{4.26}$$

where $t_s > 0$ is the sampling period. Then there exists a sufficiently small t_s such that the discrete-time dynamics (4.26) *is exponentially contracting in X, uniformly on $[-\beta, \beta]^{n_u}$.*

Proof. See the Appendix 4.A.2. □

4.4 Identification of neuronal models with the PEM

In this section, we discuss the parametric identification of conductance-based models. In Section 4.4.1, we define the identification setup, and in Section 4.4.2, we treat the case in which we can consistently identify the system's maximal conductances and reversal potentials by solving a linear least-square optimization problem.

4.4.1 Identification setup

The system (4.25) under the feedback law $i_k = \gamma(r_k - v_k)$ can be written in the form

$$x_{k+1} = x_k + t_s \left(A(u_{2,k})x_k + b(u_{2,k})\right) \tag{4.27a}$$

$$y_k = \tfrac{1}{c}\left(h(u_{2,k}, x_k) - u_{1,k} - e_k\right) \tag{4.27b}$$

and

$$v_{k+1} = v_k - t_s y_k \tag{4.28a}$$

$$u_k = \begin{bmatrix} u_{1,k} \\ u_{2,k} \end{bmatrix} = \begin{bmatrix} \gamma(r_k - v_k) \\ v_k \end{bmatrix} \tag{4.28b}$$

where x collects the states m_j and h_j for which $\alpha_j, \beta_j > 0$, $A(\cdot)$ and $b(\cdot)$ are determined by (4.9a)-(4.9b), and $g(\cdot,\cdot)$ is given by (4.11). It can be seen that by backward iteration of (4.27a) and (4.28a), we can find a F_k and a H_k such that y_k and u_k can be represented by (2.5) and (2.7), respectively. Thus, the identification setup is represented by Figure 2.11.

Assumption 5. (Input-additive white noise) The input-additive noise e_k is a sequence of independent random variables such that $E[e_k] = 0.$ and $E[e_k^2] = \sigma_e^2 > 0$. All realizations of e belong to $\mathcal{U}_\beta$.

Assumption 6. (Known signals) The signals r_k and v_k are known (v_k is measured without noise), and r is a deterministic signal that belongs to $\mathcal{U}_\beta$.

Assumption 6 indicates the fact that to identify a system with considerable additive noise at the input, an effort needs to be made to measure the output as precisely as possible. It implies that we know the signal u_k in (4.28b).

Assumption 7. For $r, e \in \mathcal{U}_\beta$, the closed-loop dynamics (4.27)-(4.28) is exponentially contracting in $V_{\beta,\gamma} \times [0,1]^{n_x}$, uniformly on $[-\beta, \beta]^2$, where $V_{\beta,\gamma}$ is given by (4.15).

It follows directly from Proposition 5 and Lemma 3 that there is a large enough value of γ, and a small enough value of t_s (possibly depending on γ) such that Assumption 7 is verified[1].

The following result will help us apply the identification framework of Section 4.2.

Proposition 6. *Under Assumptions 5-7, the closed-loop system* (4.27)-(4.28) *satisfies Assumption 1.*

Proof. See the Appendix 4.A.3. □

4.4.2 Identification with fixed ion channel kinetics

A given library of ion channel kinetic models can be used in the identification of the parameters c, $\bar{g}_j$, and ν_j of a neuronal system (recall that an ion channel kinetic model is given by the dynamics (4.9a)-(4.9b), as well as the coefficients α_j, β_j in (4.9c)). This can be done by postulating a model structure containing n_m known ion channel kinetic models, chosen a priori. For $j = 1, \ldots, n_\mathrm{m}$, we denote the objects defining the model structure kinetic models by $\hat{\tau}_{m,j}(\cdot)$, $\hat{\tau}_{h,j}(\cdot)$, $\hat{m}_{\infty,j}(\cdot)$, $\hat{h}_{\infty,j}(\cdot)$, $\hat{\alpha}_j$ and $\hat{\beta}_j$. The model structure states $\hat{m}_j$ and $\hat{h}_j$ evolve analogously to the discretized form of (4.9a)-(4.9b). These dynamics can be described in compact form by

$$\hat{x}_{k+1} = \hat{x}_k + t_s \left(\hat{A}(u_{2,k})\hat{x}_k + \hat{b}(u_{2,k}) \right) \tag{4.29}$$

where $\hat{x}$ collects the gating variables $\hat{m}_j$ and $\hat{h}_j$ for which $\hat{\alpha}_j > 0$ and $\hat{\beta}_j > 0$, respectively.

The unknown parameters of the true system, c, $\bar{g}_j$, and ν_j, all appear in the right-hand side of (4.27b), either directly or through the mapping $g(\cdot,\cdot)$ given by (4.11). To identify those parameters, we define the model structure output $\hat{y}_k(\theta)$ by

$$\begin{aligned} \hat{y}_k(\theta) = &\sum_{j=1}^{n_\mathrm{m}} (\theta_j^{(1)} + \theta_j^{(2)} u_{2,k})\, \hat{m}_{j,k}^{\hat{\alpha}_j}\, \hat{h}_{j,k}^{\hat{\beta}_j} + \\ &+ \theta_0^{(1)} + \theta_0^{(2)} u_{2,k} + \theta^{(3)} u_{1,k} \end{aligned} \tag{4.30}$$

[1]Notice that there is a tradeoff in the choice of the values of γ and t_s. Increasing the value of γ (which increases the size of $\bar{\sigma}$ in (4.47)) might require decreasing the value of t_s so that contraction of the discrete-time system is preserved.

where $u_{1,k}$ and $u_{2,k}$ are the elements of the model structure input in (4.28b), and where the vector of model structure parameters $\theta \in \mathcal{D} \subseteq \mathbb{R}^{(2n_\mathrm{m}+1)\times 1}$ has been partitioned according to

$$\theta = (\theta^{(1)\top}, \theta^{(2)\top}, \theta^{(3)})^\top$$

with $\theta^{(1)} \in \mathbb{R}^{n_\mathrm{m}\times 1}$, $\theta^{(2)} \in \mathbb{R}^{n_\mathrm{m}\times 1}$ and $\theta^{(3)} \in \mathbb{R}$.

We can more compactly write

$$\hat{y}_k(\theta) = \psi_k \theta$$

with the row vector $\psi_k \in \mathbb{R}^{1\times(2n_\mathrm{m}+1)}$ given by

$$\begin{aligned}\psi_k = \Big(&1,\ \hat{m}_{1,k}^{\hat{\alpha}_1}\,\hat{h}_{1,k}^{\hat{\beta}_1}\,,\ \ldots\ ,\ \hat{m}_{n_\mathrm{m},k}^{\hat{\alpha}_{n_\mathrm{m}}}\,\hat{h}_{n_\mathrm{m},k}^{\hat{\beta}_{n_\mathrm{m}}}\,,\ u_{2,k}\,,\\ &u_{2,k}\,\hat{m}_{1,k}^{\hat{\alpha}_1}\,\hat{h}_{1,k}^{\hat{\beta}_1},\ \ldots\ ,\ u_{2,k}\,\hat{m}_{n_\mathrm{m},k}^{\hat{\alpha}_{n_\mathrm{m}}}\,\hat{h}_{n_\mathrm{m},k}^{\hat{\beta}_{n_\mathrm{m}}},\ u_{1,k}\Big)\end{aligned}$$

Gathering ψ_k in a matrix $\Psi_N \in \mathbb{R}^{N\times(2n_\mathrm{m}+1)}$ given by

$$\Psi_N = \left[\psi_N^\top, \psi_{N-1}^\top, \ldots, \psi_1^\top\right]^\top \tag{4.31}$$

we find that the vector of model structure outputs from time $k = N$ down to time $k = 1$ is given by

$$\hat{y}_{[1,N]}^\top(\theta) = \Psi_N \theta$$

The above formulation shows that the ion channel kinetic models act as basis operators mapping the input sequence $u_{[0,N]}$, given by (4.28b), to the columns of Ψ_N. Using that formulation, we will make some assumptions in order to show that we can obtain consistent estimates of the parameter vector θ using the PEM.

First, we need all the true ion channels to be included in the model structure:

Assumption 8. The model structure contains the true system (4.27): for $j = 1, \ldots, n_\mathrm{b} \leq n_\mathrm{m}$, we have $\hat{\tau}_{m,j}(\cdot) = \tau_{m,j}(\cdot)$, $\hat{\tau}_{h,j}(\cdot) = \tau_{h,j}(\cdot)$, $\hat{m}_{\infty,j}(\cdot) = m_{\infty,j}(\cdot)$, $\hat{h}_{\infty,j}(\cdot) = h_{\infty,j}(\cdot)$, $\hat{\alpha}_j = \alpha_j$ and $\hat{\beta}_j = \beta_j$.

In case Assumption 8 is verified, then the true parameter vector, which we denote by $\bar{\theta} = (\bar{\theta}^{(1)\top}, \bar{\theta}^{(2)\top}, \bar{\theta}^{(3)})^\top$, is given by

$$\begin{aligned}
\bar{\theta}_j^{(1)} &= \begin{cases} -\bar{g}_j \nu_j / c, & j = 0, 1, \ldots, n_\mathrm{b} \\ 0, & j > n_\mathrm{b} \end{cases} \\
\bar{\theta}_j^{(2)} &= \begin{cases} \bar{g}_j / c, & j = 0, 1, \ldots, n_\mathrm{b} \\ 0, & j > n_\mathrm{b} \end{cases} \\
\bar{\theta}^{(3)} &= -1/c
\end{aligned} \tag{4.32}$$

We will also need the following standard assumption:

Assumption 9 (Persistency of excitation)**.** There is a $N^* > 0$ such that $\frac{1}{N}\Psi_N^\top \Psi_N$ and $E\left[\frac{1}{N}\Psi_N^\top \Psi_N\right]$ are positive-definite for all $N > N^*$.

Assumption 9 is an assumption both on the model structure and on r_k, the signal used to excite the true system. Intuitively, we should not include two identical ion channel kinetics in the model structure, and the excitation signal r_k should be sufficiently rich.

Finally, we make a last assumption in order to simplify our result:

Assumption 10. The initial states of the true system and of the model structure satisfy $\hat{m}_{j,0} = m_{j,0}$ and $\hat{h}_{j,0} = h_{j,0}$ for $j = 1, \ldots, n_\mathrm{b}$.

In principle, we could also estimate initial conditions. However, given the contracting nature of the ion channel kinetics considered here, initial conditions are forgotten exponentially fast. As long as Assumption 8 is verified, we can come arbitrarily close to verifying Assumption 10 by discarding initial segments of the data.

Under Assumption 6, we are able to compute

$$y_k = -\frac{v_{k+1} - v_k}{t_s}$$

from the measurements, and thus we can form the cost function $V_N(\theta)$ given by (2.10). We can now state the following result.

Theorem 1. *Let Assumptions 5-10 be satisfied. Let $N > N^*$, and let*

$$\hat{\theta}_N = (\hat{\theta}_N^{(1)\top}, \hat{\theta}_N^{(2)\top}, \hat{\theta}_N^{(3)})^\top$$

be given by

$$\begin{aligned}\hat{\theta}_N &= \arg\min_{\theta\in\mathcal{D}} V_N(\theta) \\ &= \arg\min_{\theta\in\mathcal{D}} \frac{1}{N}\|y_{[1,N]}^\top - \Psi_N\theta\|_2^2\end{aligned} \tag{4.33}$$

where y_k and Ψ_N are given by (4.27b) *and* (4.31)*, respectively, and $\mathcal{D}$ is a compact parameter domain containing $\bar{\theta}$, the true parameter vector* (4.32)*. Then, we have $\hat{\theta}_N \to \bar{\theta}$ w.p. 1 as $N \to \infty$.*

Proof. By Assumptions 8 and 10, the true output y_k, given by (4.27b) (see also (4.11)), can be written as

$$y_{[1,N]}^\top = \Psi_N\bar{\theta} - \tfrac{1}{c}e_{[1,N]}^\top$$

and thus we can write

$$E\left[V_N(\theta)\right] = \tfrac{1}{N}E\left[\|\Psi_N(\bar{\theta}-\theta) - \tfrac{1}{c}e_{[1,N]}^\top\|_2^2\right]$$

By Assumption 5, the time-delay present in the system ensures that v_k and x_k do not depend on e_k. We then have that

$$E\left[\Psi_N^\top e_{[1,N]}^\top\right] = 0$$

and thus

$$E\left[V_N(\theta)\right] = \tfrac{1}{N}(\bar{\theta}-\theta)^\top E\left[\Psi_N^\top\Psi_N\right](\bar{\theta}-\theta) + \tfrac{1}{c^2}\sigma_e^2$$

Using Assumption 9, we have

$$\arg\min_{\theta\in\mathcal{D}} E\left[V_N(\theta)\right] = \bar{\theta} \tag{4.34}$$

for all $N > N^*$.

It remains to show that $\hat{\theta}_N$ converges to (4.34) w.p. 1 as $N \to \infty$, which can be done by means of Lemma 1. Assumption 1 is satisfied due to Proposition 6. By Assumptions 6 and 7, the input (4.28b) to the model structure belongs to $\mathcal{U}_{\beta^*}^2$ for some $\beta^* > 0$, since v_k remains in the compact set $V_{\beta,\gamma}$. Proposition 4 and Lemma 2 then ensure that the model structure verifies Assumption 2. It follows by Lemma 1 that $V_N(\theta)$ converges uniformly to $E\left[V_N(\theta)\right]$ on the compact set $\mathcal{D}$. In view of (4.33) and (4.34), this ensures the result of the theorem. □

4.5 Examples

In this section, we illustrate the results of Section 4.4.2 by identifying various discrete-time neuronal models. All discrete-time models are obtained by forward-Euler discretization of their continuous-time counterparts with $t_s = 0.005$ ms.

Example 5. In this example, we identify the discretized Hodgkin-Huxley (HH) model of Example 3. We include in the model structure the two ion channel kinetics present in the true model, and identify the values of c, $\bar{g}_j$, and ν_j using the parameter vector θ. From (4.12) and (4.30), we have the following true parameters:

$\bar{\theta}_0^{(1)}$	$\bar{\theta}_0^{(2)}$	$\bar{\theta}_1^{(1)}$	$\bar{\theta}_1^{(2)}$	$\bar{\theta}_2^{(1)}$	$\bar{\theta}_2^{(2)}$	$\bar{\theta}^{(3)}$
$0.3 \cdot 54.4$	0.3	$120 \cdot -55$	120	$36 \cdot 77$	36	-1

We simulated an identification experiment in which the discretized HH model is subject to output feedback with $\gamma = 50$. The model is excited by the user-defined input $r_k = -45 + \tilde{r}_k$, where $\tilde{r}_k$ is given by white Gaussian noise of standard deviation $\sigma_r = 100$ that is filtered by the zero-order hold discretization of the system $10^2/(s+10)^2$. The input noise e_k was generated with $\sigma_e = 2.5$ and was truncated so that $e_k = 100$ whenever $e_k \geq 100$. This setup resulted in a signal-to-noise ratio (between y_k and e_k) of around 30.8 dB. To eliminate transient effects and satisfy Assumption 10 as close as possible, we eliminated the initial 0.5 seconds of measurement (corresponding to 10^5 samples) from all datasets.

Figure 4.5 shows the estimation error $\bar{\theta} - \hat{\theta}_N$ for $N = 1 \times 10^5$ to $N = 9 \times 10^5$ (corresponding to experiment times of 0.5 to 4.5 seconds) for 20 different realizations of the experiment, as well as their average. We can see from the figure that the estimates steadily converge to the true parameters.

Example 6. In this example, we illustrate how a library of pre-established set of ion channel kinetic models can be used to identify different neuronal models. We consider three models, all of which are based on a forward-Euler discretization of the system given by

$$\dot{v}(t) = -0.3(v(t) + 17) - \sum_{j=1}^{4} i_j(t) + i(t) + e(t)$$

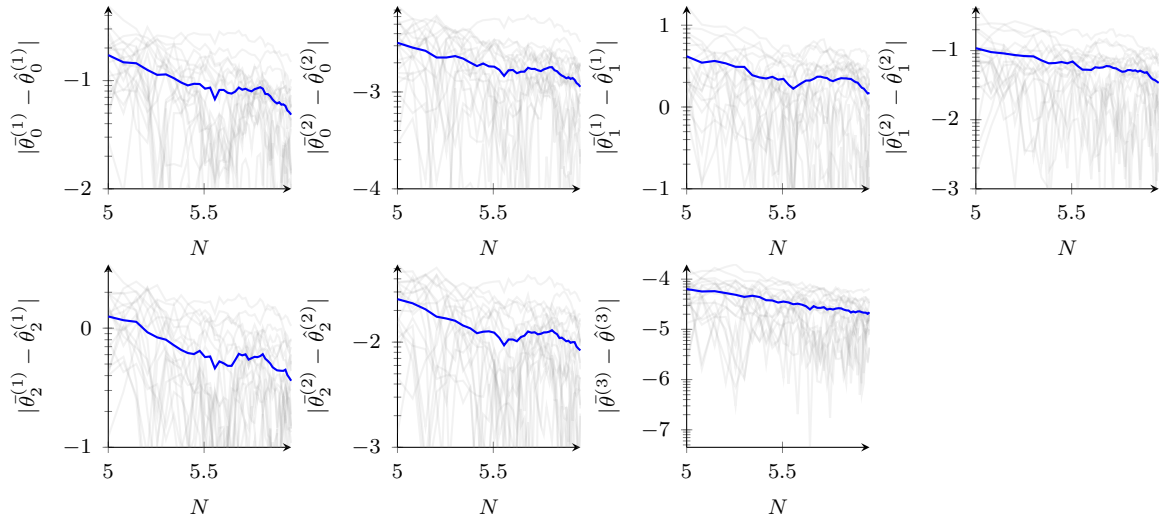


Fig. 4.5 The $\log_{10} \times \log_{10}$ plots above show how the errors in the estimated parameters of Example 5 fall as the number of data points N increases. In grey: errors in each of the 20 realizations of the identification experiment as computed for $N = 1 \times 10^5$ to 9×10^5 ($t_s = 0.005$). In blue: average of the 20 error traces.

with

$$
\begin{aligned}
i_1 &= 120\, m_1^3\, h_1(v - 55) \\
i_2 &= 20\, m_2^4 (v + 75) \\
i_3 &= \bar{g}_3\, m_3^3\, h_3(v + 75) \\
i_4 &= \bar{g}_4\, m_4^2 (v - 120)
\end{aligned}
$$

where the states m_j and h_j are given by (4.9a) and (4.9b), respectively with the functions $m_{\infty,j}$, $h_{\infty,j}$, $\tau_{m,j}$ and $\tau_{h,j}$ plotted in Figure 4.6 and described in Appendix 4.C.

The above system, taken from [31], defines a modified version of the Connor-Stevens neuronal model [23]. The values of the variables $\bar{g}_3$ and $\bar{g}_4$ are the distinguishing factors between the three models we use in this example. We call them Connor-Stevens (CS) models A, B, and C, according to the following maximal conductance values:

CS model	A	B	C
$\bar{g}_3$	0	90	0
$\bar{g}_4$	0	0	0.4

Connor Stevens model A is similar to the HH model of the previous example, while models B and C differ from A due to the addition of ion currents i_3 and i_4,

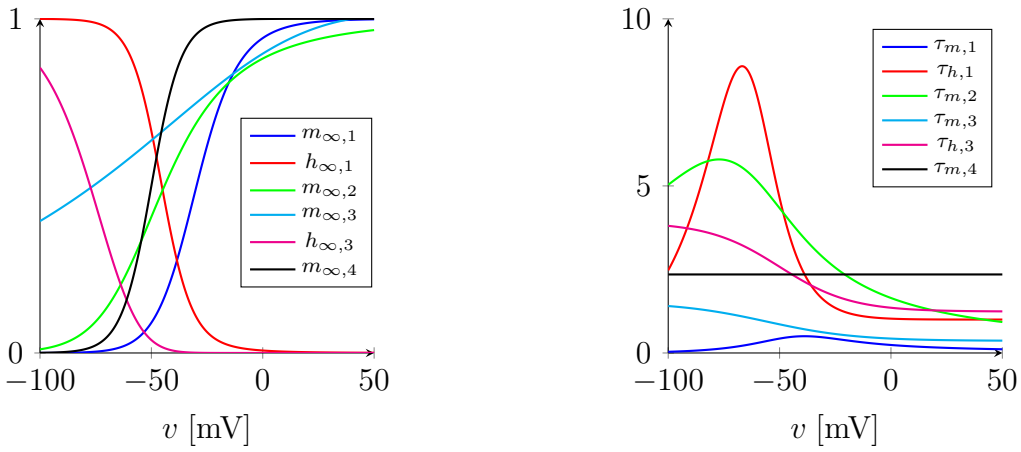


Fig. 4.6 Left: Nonlinear activation functions $m_{\infty,j}$ and $h_{\infty,j}$ in the Connor-Stevens model. Right: Time constant functions $\tau_{m,j}$ and $\tau_{h,j}$ in the Connor-Stevens model.

respectively (these currents represent an "A-type" potassium current and a calcium current, respectively). It can be verified through simulations that the addition of i_3 or i_4 makes the qualitative input-output behavior (from i to v) of models B and C differ from that of model A. In particular, models B and C can fire periodic spikes with arbitrarily low frequency, while model A does not have that property (see, for instance, Figure 2 of [31]). The property of spiking with arbitrarily low frequency has important neurocomputational consequences. It underlies the classical distinction between Type I and Type II neuronal excitability first proposed by Hodgkin and Huxley (see [63, Chapter 7]).

To identify the models A, B and C, we include in a single model structure all four of the ion channels shared by those models. We simulated an identification experiment in which the discretized CS models are subject to output feedback with $\gamma = 50$. All models are excited by the user-defined input $r_k = -45 + \tilde{r}_k$, where $\tilde{r}_k$ is given by white Gaussian noise of standard deviation $\sigma_r = 30$ that is filtered by the zero-order hold discretization of the system $10^2/(s+10)^2$. The input noise e_k was generated with $\sigma_e = 1$ and was truncated so that $e_k = 100$ whenever $e_k \geq 100$. This setup resulted in a signal-to-noise ratio (between y_k and e_k) of around 28 dB, 26 dB and 29 dB for the CS models A, B and C, respectively (again, we eliminated the first 0.5 seconds of measurement from all datasets).

Figure 4.7 shows the evolution of the estimates of $\bar{g}_j$ obtained by identifying each of the CS models A, B and C (for brevity, we do not show the evolution of all parameter estimates). It can be seen that the estimates of $\bar{g}_3$ (or $\bar{g}_4$) for models that do not

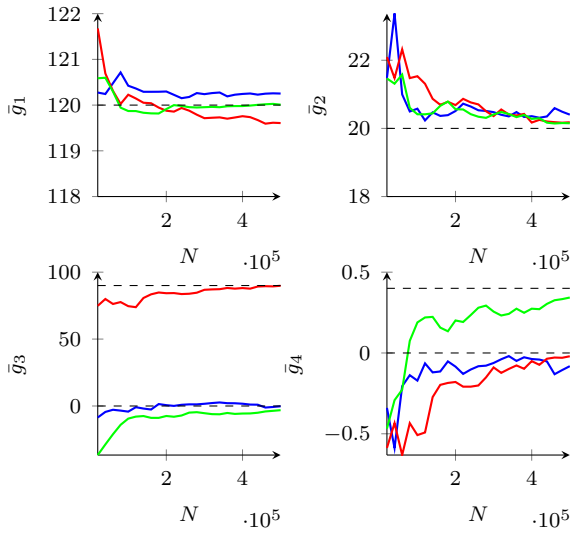


Fig. 4.7 Evolution of the estimates of $\bar{g}_j$, with respect to the number of samples, for each of the identified Connor-Stevens models A (blue), B (red) and C (green). The sampling period is $t_s = 0.005$, and the experimental setup is described in Example 6.

contain i_3 (or i_4) tend towards zero, while the other estimates tend towards their true values.

4.6 Model structure discrimination

In this section, inspired by the work in [127], we show that the PEM framework can be used in a pre-processing stage to aid in choosing a model structure for the internal dynamics of a neuronal model. The idea is simple: by using output feedback, we can stabilize the system (4.25) around a unique arbitrary equilibrium point. The behavior of the nonlinear system close to this arbitrary equilibrium is approximately that of the linearized system. The pattern of movement of the poles and zeros of the linearized system as we change the system equilibrium can provide valuable information about the structure of the internal dynamics. We can thus use a linear model structure in conjunction with small excitation signals to estimate and analyze these linearizations.

4.6.1 Linearized internal dynamics

We start with a simple fact about contracting systems.

Proposition 7. *Assume the discrete-time dynamics* (4.1a) *is contracting, with a constant metric P, and that X is convex, closed and bounded. Assume that the input*

is given by a constant, $u_k = \bar{u}$ for all $k \geq 0$. Then (4.1a) *has a unique globally exponentially stable equilibrium point.*

Proof. The mapping $f(x, \bar{u})$ is a continuous mapping from the convex compact set $X \subset \mathbb{R}^n_x$ to itself, and thus it has a fixed point. It follows from (4.45), with $u_k = \tilde{u}_k = \bar{u}$, that the fixed point is unique and exponentially stable. □

We will consider the more general dynamics

$$v_{k+1} = v_k + \tfrac{t_s}{c}\left(-h(v_k, x_k) + i_k + e_k\right) \tag{4.35a}$$

$$x_{k+1} = x_k + t_s f(v_k, x_k), \tag{4.35b}$$

and assume t_s is small enough so that the internal dynamics (4.35b) is exponentially contracting. By Proposition 7, the equation

$$0 = f(v^{\text{eq}}, x^{\text{eq}})$$

has a unique solution x^{eq}. We can use the solution to define a mapping $x_\infty : V \to X$ such that $x^{\text{eq}} = x_\infty(v^{\text{eq}})$ and $f(v^{\text{eq}}, x_\infty(v^{\text{eq}})) = 0$. This can in turn be used to define a mapping $i_\infty : V \to Y$ (the static I-V curve) through

$$i_\infty(v) = h(v, x_\infty(v)) \tag{4.36}$$

From (4.35a), with $e_k = 0$, this is the steady-state input current i associated with the equilibrium v.

Let $\delta v_k = v_k - v^{\text{eq}}$ and $\delta y_k = y_k - i_\infty(v^{\text{eq}})$, $k \geq 0$, and let $\Delta V(z) = \mathcal{Z}[\delta v_k]$ and $\Delta Y(z) = \mathcal{Z}[\delta y_k]$, where $\mathcal{Z}$ denotes the z-transform. The transfer function $G_{\text{int}}(z; v^{\text{eq}})$ such that $\Delta Y(z) = G_{\text{int}}(z; v^{\text{eq}})\Delta V(z)$ is given by

$$G_{\text{int}}(z; v^{\text{eq}}) = \frac{\partial h}{\partial x}\left[zI - \left(I + t_s \frac{\partial f}{\partial x}\right)\right]^{-1} t_s \frac{\partial f}{\partial v} + \frac{\partial h}{\partial v}$$

where all partial derivatives are evaluated at v^{eq} and $x_\infty(v^{\text{eq}})$.

By analyzing the pattern of movement of poles and zeros of $G_{\text{int}}(z; v^{\text{eq}})$, as we change v^{eq}, it is possible to obtain information about the model structure of the internal dynamics (4.35b). If the zeros of $G_{\text{int}}(z; v^{\text{eq}})$ move, but the poles do not move, we have an indication that the *zero dynamics* $x_{k+1} = x_k + t_s f(x_k, 0)$ is linear. In this case, the internal system could be modeled by that of a Parallel-Wiener, Parallel-Hammerstein, or Parallel-Wiener-Hammerstein system [129]. If the zeros *and* poles of $G_{\text{int}}(z; v^{\text{eq}})$

move with v^{eq}, then the internal system could be modeled by a Hodgkin-Huxley type conductance-based model, or by polynomial nonlinear state-space models such as those used in [100], or by a universal approximator such as the one introduced in the next chapter.

To find out what the zero-pole movement pattern is, we estimate linear approximations of the system at different setpoints, as detailed in the next section.

4.6.2 Approximation of the linearized dynamics

Consider the closed-loop identification setup given by

$$x_{k+1} = x_k + t_s f\left(u_k, x_k\right) \tag{4.37a}$$

$$y_k = h\left(u_k, x_k\right) - e_k \tag{4.37b}$$

and

$$v_{k+1} = v_k + \tfrac{t_s}{c}\Big(- y_k + \gamma(r_k - v_k)\Big) \tag{4.38a}$$

$$u_k = v_k \tag{4.38b}$$

where the feedback block (4.38) is known. This is similar to the setup (4.27), except that now we assume knowledge of c, and the dynamics of the system is more general.

Consider an external input $r_k = \bar{r} + \tilde{r}_k$, where $\bar{r} \in \mathbb{R}$ is a constant, and assume that the system dynamics is exponentially contracting. Let $v_\infty(\bar{r})$ denote the steady-state voltage of the system subject to the input $r_k = \bar{r}$. Note that the same arguments of Section 4.6.1 can be used to show that v_∞ is well-defined.

We use a linear model to find a linear approximation for $G_{\text{int}}(z; v_\infty(\bar{r}))$. The model structure is given by

$$\hat{y}_k(\theta) = \hat{G}_{\text{lin}}(q;\theta)(v_k - v_\infty(\bar{r})) + i_\infty(v_\infty(\bar{r})) \tag{4.39}$$

where q denotes the forward shift operator. In this model structure, we assume that the mappings v_∞ and i_∞ are known. Although they could both be estimated using steady-state static data (as in Chapter 3), in practice it may be simpler to take

$$\begin{aligned} v_\infty(\bar{r}) &\approx \frac{1}{N}\sum_{k=1}^{N} v_k \\ i_\infty(v_\infty(\bar{r})) &\approx \frac{1}{N}\sum_{k=1}^{N} y_k \end{aligned} \tag{4.40}$$

since the time-varying part of $r_k = \bar{r} + \tilde{r}_k$ is meant to be small in this experiment.

Assumption 11. The transfer function $\hat{G}_{\text{lin}}(z;\theta)$ is parameterized according to

$$\hat{G}_{\text{lin}}(z;\theta) = \frac{b_0 + b_1 z^{-1} + \ldots + b_{n_b} z^{-n_b}}{1 + a_1 z^{-1} + \ldots + a_{n_a} z^{-n_a}}$$

where the variables b_i and a_i are collected in the vector θ. The parameter domain $\mathcal{D}$ is given by the set of θ such that the magnitude of the poles of $\hat{G}_{\text{lin}}(z;\theta)$ are bounded by a number smaller than one.

Assumption 12. $\bar{V}(\theta) \triangleq \lim_{N\to\infty} E\left[V_N(\theta)\right]$ exists.

A sufficient condition for Assumption 12 to hold is that the prediction error

$$\epsilon(k,\theta) = (y_k - \hat{y}_k(\theta))^2$$

is quasi-stationary (see [77, Definition 2.1]). For stable linear systems, a quasi-stationary input produces a quasi-stationary output [77, Theorem 2.2]. Thus, as long as $\tilde{r}_k + e_k$ is a quasi-stationary signal of small amplitude (so that the behavior of the nonlinear system is close to linear), the assumption above is not unreasonable[2].

Theorem 2. *Let Assumptions 5-6 and 11-12 be satisfied. Assume the closed-loop dynamics* (4.37)-(4.38) *is exponentially contracting in a convex, closed and bounded domain of the state-space, uniformly on* $U = [-\beta, \beta]^{n_u}$. *Let* $\hat{\theta}_N$ *be given by* (2.11), *and* $\hat{y}_k(\theta)$ *be given by* (4.39). *Then* $\hat{\theta}_N \to \arg\min_{\theta\in\mathcal{D}} \bar{V}(\theta)$.

Proof. By Assumption 11, (4.39) and its derivatives with respect to a_i and b_i are stable for all values of θ; thus Assumption 2 is satisfied. To ensure Assumption 1, we can proceed similarly to the proof of Proposition 6. Convergence of the parameter estimates is then guaranteed by Lemma 1 and Assumption 12. □

Suppose now that we define a sequence of values for $\bar{r}$. For every value in the sequence, we solve the estimation problem $\min_\theta V_N(\theta)$ and obtain a parameter estimate $\hat{\theta}_N(\bar{r})$. We can thus attempt to infer the poles of G_{int} from those of $\hat{G}_{\text{lin}}(z;\hat{\theta}_N(\bar{r}))$. The next example illustrates this approach, including the effects that the signal-to-noise ratio will have on it.

[2]The error $\epsilon(k,\theta)$ will be "close" to quasi-stationarity, with "close" meaning that $\limsup_{N\to\infty} E[V_N(\theta)]$ is close to $\liminf_{N\to\infty} E[V_N(\theta)]$. For a rigorous discussion on these limits and how they relate to the parameter estimates, see [76].

Example 7. In this example, we attempt to identify the poles of the linearized inverse of the Hodgkin-Huxley model. To generate data for this example, we simulated a discretized Hodgkin-Huxley model under output feedback, with $t_s = 10^{-2}$ and $\gamma = 100$.

We performed multiple experiments using the excitation signal $r_k = \bar{r} + \tilde{r}_k$, where

$$\bar{r} \in \{-80, -70, -60, \ldots, 10\} \quad \text{mV}$$

and $\tilde{r}_k$ is white Gaussian noise of standard deviation $\sigma_r = 2$ mV that is saturated for $|\tilde{r}_k| \geq 100$ mV. The input noise e_k is Gaussian white noise with $\sigma_e = 1$ μA/cm^2 that is saturated in the same way as $\tilde{r}_k$. We simulated 2 seconds of experiment (corresponding to 2×10^5 samples), removing the first 0.1 seconds from the datasets.

We choose the orders $n_a = 3$ and $n_b = 3$ for the transfer function parametrization of Assumption 11, so as to match the order of the nonlinear system. To compute the cost function $V_N(\theta)$, we used the approximation (4.40). The model structure leads to a nonlinear least-squares optimization problem $\min_\theta V_N(\theta)$. To solve it, we use the Levenberg-Marquardt nonlinear search method[3].

In Figure 4.8 we show the time-constant functions of the continuous-time Hodgkin-Huxley model (Example 3) and the poles of the linear systems that have been estimated according to the procedure described above. The discrete-time identified poles p_{DT} have been converted to continuous-time poles p_{CT} by means of the transformation $p_{CT} = (p_{DT} - 1)/t_s$, and the continuous-time time constants are given by $-1/p_{CT}$.

It can be seen in Figure 4.8 that, in most experiments, two of the identified poles match actual poles of the Hodgkin-Huxley model. These identified poles provide evidence that a model structure with nonlinear internal dynamics is necessary for the nonlinear identification of the system. In case a classical conductance-based model structure is chosen, the poles can be used to construct the τ functions of the model by fitting a continuous function to their values. It is interesting to note that the procedure fails to identify the third moving pole of the model (related to $\tau^{(1,2)}$). In its place, the identified model contains a spurious very fast pole. The reason for this failure becomes clear when we plot the ground truth pole $-1/\tau^{(1,2)}(v)$ against one of the three zeros obtained through exact linearization of G_{int}. This plot is shown in Figure 4.9. We can see that there is a near pole-zero cancellation, resulting in the pole (and the zero) not being identified.

In Figure 4.10, we display the signal-to-noise ratio (SNR) of y_k relative to e_k for each experiment, as well as the normalized root mean squared error (NRMSE) for each

[3]This was implemented using Matlab's `oe` routine with the options `'SearchMethod'` set to `'lm'` and `'InitialCondition'` set to `'zero'`.

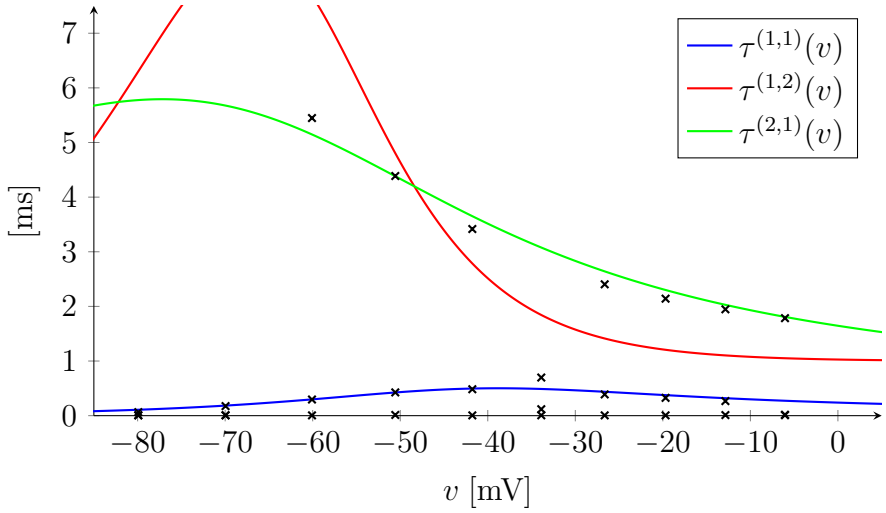


Fig. 4.8 Time-constants of the linear systems identified around the mean outputs $\frac{1}{N}\sum_{k=1}^{N} v_k$ due to small-variation inputs $r_k = \bar{r} + \tilde{r}_k$. The discrete-time time-constants have been converted to continuous-time (following the forward-Euler rule) for comparison with the Hodgkin-Huxley τ functions. Each column of points corresponds to the poles of one third-order linear system. Two of the ground-truth poles are identified in most experiments. The failure to identify the last pole can be attributed to a zero-pole cancellation in the ground truth linearized model (see Figure 4.9).

experiment. The signal-to-noise ratio of the experiments vary significantly, and this affects the quality of the identified linear models. The variation in signal-to-noise ratio can be attributed to the fact that the IV curve $i_\infty(\bar{r})$ of the ground-truth system varies significantly over the range of setpoints $\{-80, \ldots, 10\}$, whereas the noise level of e_k is the same for all experiments. The two first experiments (corresponding to $\bar{r} = -80$ and $\bar{r} = -70$) fail to properly identify any of the slow time constants (see Figure 4.8) due to the low signal-to-noise ratio. In that case, more data needs to be collected.

To solve this SNR issue, one solution could be to increase the variance of $\tilde{r}$. However, as this variance increases, the effect of the system nonlinearity in the obtained data also increases. Thus, there is a tradeoff between keeping the behavior of the system close to linear, and maintaining a good SNR.

In Example 7, the failure to identify one of the ground-truth poles of the Hodgkin-Huxley model (due to a zero-pole cancellation) hints that this model could be amenable to reduction: for instance, we could use a model structure that *omits* the state $x^{(1,2)}$ related to the unidentified pole. This state is related to the *inactivation* of the sodium ion channel, which provides a slow negative feedback to the membrane current – the same role provided by *activation* of the Potassium current related to the state

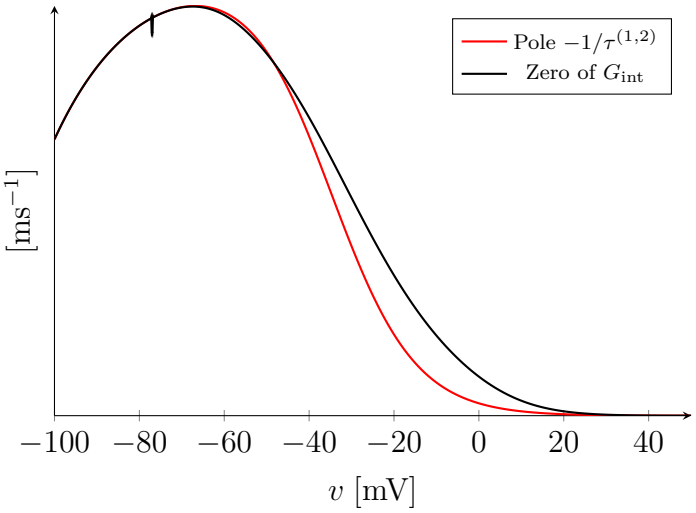


Fig. 4.9 Near pole-zero cancellation in the linearized Hodgkin-Huxley model: the pole $-1/\tau^{(1,2)}$ of G_{int} is nearly canceled by one of the zeros of G_{int}.

$x^{(2,1)}$. Going a step further and omitting the state $x^{(1,1)}$ (which has a very small time constant), we recover the model structure of a classical two-state reduced version of the Hodgkin-Huxley neuron [66, Section 5.1.3].

4.7 Conclusion

In this chapter, we studied the identification of neuronal systems under the assumption of current-additive zero-mean white noise and negligible voltage measurement noise. We showed that by treating a neuronal model as a closed-loop system, we can solve the identification problem by identifying the inverse dynamics with an output-error model structure. We have demonstrated that consistent parameter estimates are obtained when the model structure contains the internal dynamics of the system being identified. As seen in Section 2.4.1, this is a common strategy adopted in neuroscience, where kinetic models of ion channels are estimated in separate experiments (see also [90]). It is worth noting that the results in this chapter may hold for ion channel models which are more general than (4.9a)-(4.9b); the key requirement is that the ion channels possess a contracting dynamics, so that (4.13) is satisfied. Thus, this work rigorously justifies the estimation of conductance-based models using conventional methods of nonlinear identification.

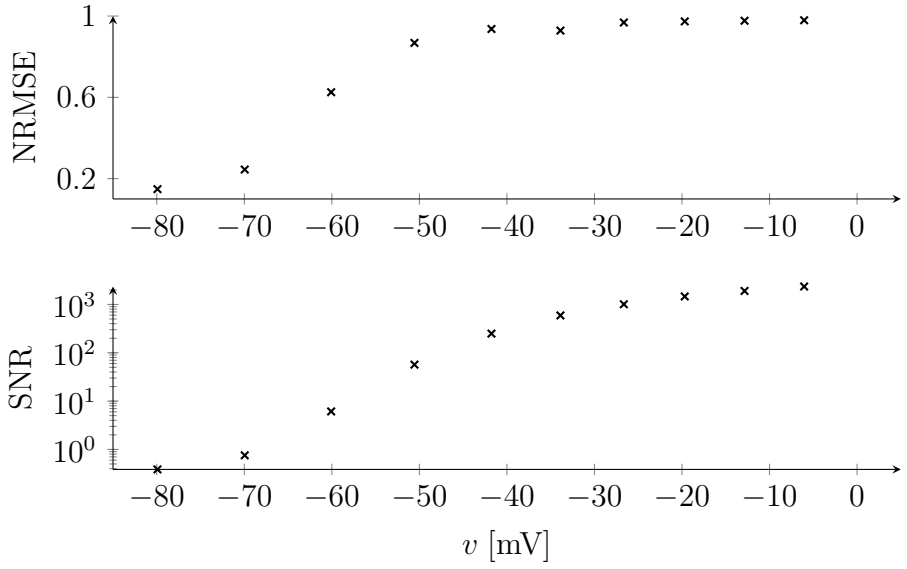


Fig. 4.10 Normalized root mean squared error (NRMSE) and Signal-to-noise ratio (SNR) measured at each identification experiment. Due to the input-additive noise assumption, the noise remains the same in all experiments, and the signal-to-noise ratio follows the shape of the Hodgkin-Huxley IV curve $i_\infty(v)$. All parameters are described in Example 7.

Appendix

4.A Proofs

4.A.1 Proof of Lemma 2

Let $P = \Theta^\top \Theta$ ($\Theta > 0$). For $k \geq 0$, let the sequences x_k and $\tilde{x}_k$ be defined by the dynamics (4.1a) subject to the inputs u_k and $\tilde{u}_k$, respectively. Applying the change of coordinates $z_k = \Theta x_k$, we obtain the discrete-time dynamics

$$z_{k+1} = f_\Theta(z_k, u_k) \tag{4.41}$$

with f_Θ given by

$$f_\Theta(\zeta, \upsilon) = \Theta f\left(\Theta^{-1}\zeta, \upsilon\right), \tag{4.42}$$

The inequality (4.2) implies that

$$\sigma_{\max}\left[\Theta \frac{\partial f}{\partial x}(x_k, u_k)\Theta^{-1}\right] \leq \alpha < 1 \tag{4.43}$$

for all $x_k \in X$ and $k \in \mathbb{Z}_+$. The inequality (4.43) in turn implies that the induced 2-norm of $\partial f_\Theta / \partial \zeta$ satisfies $\|\partial f_\Theta / \partial \zeta\|_2 \leq \alpha$ whenever

$$\zeta \in Z = \{\zeta \in \mathbb{R}^n_x \mid \zeta = \Theta\xi,\ \xi \in X\}$$

where (by convexity of X) Z is a convex set. Furthermore, since $\partial f_\Theta / \partial \upsilon$ is a continuous function and $Z \times [-\beta, \beta]^{n_u}$ is closed and bounded, there is some $L_1 > 0$ such that $\|\partial f_\Theta / \partial \upsilon\|_2 \leq L_1$ for all $(\zeta, \upsilon) \in Z \times [-\beta, \beta]^{n_u}$.

Now let $\zeta, \tilde{\zeta} \in Z$ and $\upsilon, \tilde{\upsilon} \in [-\beta, \beta]^{n_u}$. Let also $\gamma_1(s) = (1-s)\tilde{\zeta} + s\zeta$ and $\gamma_2(s) = (1-s)\tilde{\upsilon} + s\upsilon$, with $s \in [0, 1]$. It can be shown, using the mean value theorem

(see, e.g., the proof of [67, Lemma 3.1]), that there is a $s^* \in (0,1)$ such that

$$\|f_\Theta(\zeta,\upsilon) - f_\Theta(\tilde{\zeta},\tilde{\upsilon})\|_2 \leq \left\| \frac{\partial f_\Theta}{\partial \zeta}(\gamma_1(s^*),\gamma_2(s^*))(\zeta-\tilde{\zeta}) + \frac{\partial f_\Theta}{\partial \upsilon}(\gamma_1(s^*),\gamma_2(s^*))(\upsilon-\tilde{\upsilon}) \right\|_2$$

By convexity of $Z \times [-\beta,\beta]^{n_u}$, the above implies

$$\|f_\Theta(\zeta,\upsilon) - f_\Theta(\tilde{\zeta},\tilde{\upsilon})\|_2 \leq \alpha\|\zeta-\tilde{\zeta}\|_2 + L_1\|\upsilon-\tilde{\upsilon}\|_2 \tag{4.44}$$

By recursive application of (4.41) and (4.44), we have

$$\|z_k - \tilde{z}_k\|_2 \leq L_1 \sum_{m=1}^{k} \alpha^{m-1}\|u_{k-m} - \tilde{u}_{k-m}\|_2 + \alpha^k\|z_0 - \tilde{z}_0\|_2$$

for $k \geq 0$. Multiplying both sides of the inequality by $\|\Theta^{-1}\|_2$ and substituting $z_k = \Theta x_k$, we have

$$\begin{aligned}\|x_k - \tilde{x}_k\|_2 \leq & \frac{L_1}{\sigma_{\min}} \sum_{m=1}^{k} \alpha^{m-1}\|u_{k-m} - \tilde{u}_{k-m}\|_2 \\ & + \frac{\sigma_{\max}}{\sigma_{\min}}\, \alpha^k\, \|x_0 - \tilde{x}_0\|_2\end{aligned} \tag{4.45}$$

for $k \geq 0$, where $\sigma_{\max}$ and $\sigma_{\min}$ denote the largest and the smallest singular values of Θ, respectively.

By arguments similar to those used above, there are $L_2, L_3 > 0$ such that

$$\|y_k - \tilde{y}_k\|_2 \leq L_2\|x_k - \tilde{x}_k\|_2 + L_3\|u_k - \tilde{u}_k\|_2 \tag{4.46}$$

The result (4.4) follows directly from (4.46) and (4.45) by setting $C_1 = \max\{L_1 L_2 \sigma_{\min}^{-1}, L_3\}$ and $C_2 = L_2\sigma_{\max}\sigma_{\min}^{-1}$.

4.A.2 Proof of Lemma 3

Since $\partial f/\partial x$ is continuous on the closed and bounded set $X \times [-\beta,\beta]^{n_u}$, there is a number $\bar{\sigma}$ such that $\bar{\sigma} \geq \sigma_{\max}[\partial f/\partial x]$ on that set. Let $f_{t_s}(x_k,u_k) = x_k + t_s f(x_k,u_k)$.

Using (4.6), we have

$$\begin{aligned}\frac{\partial f_{t_s}^\top}{\partial x_k} P \frac{\partial f_{t_s}}{\partial x_k} &= \left(1 + t_s \frac{\partial f^\top}{\partial x_k}\right) P \left(1 + t_s \frac{\partial f}{\partial x_k}\right) \\ &\leq (1 - 2t_s\lambda)P + t_s^2 \frac{\partial f^\top}{\partial x_k} P \frac{\partial f}{\partial x_k} \\ &\leq \left(1 - 2t_s\lambda + t_s^2 \frac{\lambda_{\max}[P]}{\lambda_{\min}[P]} \bar{\sigma}^2\right) P \\ &= \alpha(t_s)^2 P\end{aligned} \tag{4.47}$$

for all $(x_k, u_k) \in X \times [-\beta, \beta]^{n_u}$. The second inequality above is obtained using the fact that $A^\top P A \leq \lambda_{\max}[P]\sigma_{\max}^2[A]I$ and $I \leq 1/\lambda_{\min}[P]P$. Making $t_s < 1$ small enough ensures that $\alpha(t_s)^2 < 1$, and thus that the dynamics (4.26) is exponentially contracting on X, uniformly on $[-\beta, \beta]^{n_u}$.

4.A.3 Proof of Proposition 6

Consider two different solutions of (4.27)-(4.28). The first is given by

$$\begin{aligned}(v_k, x_k^\top)^\top &= \phi_{k,0}(\gamma r + e, (v_0, x_0^\top)^\top) \\ (v_0, x_0^\top)^\top &= 0\end{aligned} \tag{4.48}$$

for $k \geq 0$, and the second is given by

$$\begin{aligned}(\bar{v}_{k,s+1}, \bar{x}_{k,s+1}^\top)^\top &= \phi_{k,s+1}(\gamma r + e, (\bar{v}_{s+1}, \bar{x}_{s+1}^\top)^\top) \\ (\bar{v}_{s+1}, \bar{x}_{s+1}^\top)^\top &= 0\end{aligned} \tag{4.49}$$

for $k \geq s + 1$.

We will use the solutions above to construct the random variables $\bar{y}_{k,s}$ involved in Assumption 1. First, for each $s \in \mathbb{Z}_+$, we set $\bar{y}_{s,s} = 0$. From (4.27b), we compute the sequence y_k using (4.48), for $k \geq 0$, and the sequence $\bar{y}_{k,s+1}$ using (4.49), for $k \geq s + 1$. We have that $\bar{y}_{k,s}$ is independent of $e_{[0,s]}$, since $e_{[s+1,k]}$ is independent of $e_{[0,s]}$; furthermore, r_k is deterministic; thus the independence required in Assumption 1 is satisfied. We now need to verify (2.12a) for $k \geq s$. For $k = s$, we have

$$\begin{aligned}|y_s - \bar{y}_{s,s}| = |y_s| &= \tfrac{1}{c}\big|h(v_s, x_s) - \gamma(r_s - v_s) - e_s\big| \\ &\leq \tfrac{1}{c}\Big(|h(v_s, x_s) + \gamma v_s| + (\gamma + 1)\beta\Big) \\ &\leq C_1\end{aligned} \tag{4.50}$$

for some $C_1 > 0$ and for each $s \in \mathbb{Z}_+$. To ensure this bound, we have used (from Assumptions 5 and 6) the fact that $r, e \in \mathcal{U}_\beta$, and the fact that $h(v,x) + \gamma v$ is a continuous function on the compact $V_{\beta,\gamma} \times [0,1]^{n_x}$.

Now, we make use of Assumption 7. Let $\alpha_c < 1$ be the contraction rate of the closed-loop dynamics (4.27)-(4.28). Since $|\gamma r + e| \leq (\gamma+1)\beta$, we can apply Lemma 2 (with the time origin shifted to $s+1$) to see that there is a $C_2 > 0$ such that

$$\begin{aligned} |y_k - \bar{y}_{k,s+1}| &\leq C_2 \alpha_c^{k-(s+1)} \, \|(v_{s+1}, x_{s+1}^\top) - (\bar{v}_{s+1}, \bar{x}_{s+1}^\top)\| \\ &= \alpha_c^{-1} C_2 \, \alpha_c^{k-s} \, \|(v_{s+1}, x_{s+1}^\top)\| \\ &\leq \alpha_c^{-1} C_2 \, C_3 \, \alpha_c^{k-s} \end{aligned} \tag{4.51}$$

for each $s \in \mathbb{Z}_+$ and $k \geq s+1$, where the constant $C_3 > 0$ comes from the boundedness of $V_{\beta,\gamma} \times [0,1]^{n_x}$. Taking $E[\,\cdot^4]$ on both sides of (4.50) and (4.51), we verify (2.12a) with $C = \max\{C_1^4, (\alpha_c^{-1} C_2 C_3)^4\}$ and $\alpha = \alpha_c^4$.

The random variables $\bar{u}_{k,s}$ of Assumption 1 can be constructed in a completely analogous way, and thus we omit this part of the proof.

4.B Hodgkin-Huxley kinetic functions

To define the ion channel kinetics of the Hodgkin-Huxley model, we first set

$$\begin{aligned} \alpha_{m,1}(v) &= 0.1 \frac{-40 - v}{\exp\left(\frac{-40-v}{10}\right) - 1} & \beta_{m,1}(v) &= 4 \exp\left(\tfrac{-v-65}{18}\right) \\ \alpha_{h,1}(v) &= 0.07 \exp\left(\tfrac{-v-65}{20}\right) & \beta_{h,1}(v) &= \frac{1}{\exp\left(\frac{-35-v}{10}\right) + 1} \\ \alpha_{m,2}(v) &= 0.01 \frac{-55 - v}{\exp\left(\frac{-55-v}{10}\right) - 1} & \beta_{m,2}(v) &= 0.125 \exp\left(\tfrac{-v-65}{80}\right) \end{aligned}$$

Then, the functions $\tau_{m,j}$ and $m_{\infty,j}$, $j = 1, 2$, are given by

$$\begin{aligned} \tau_{m,j}(v) &= \frac{1}{\alpha_{m,j}(v) + \beta_{m,j}(v)} \\ m_{\infty,j}(v) &= \frac{\alpha_{m,j}(v)}{\alpha_{m,j}(v) + \beta_{m,j}(v)} \end{aligned} \tag{4.52}$$

The same relationships are used to define $\tau_{h,1}$ and $h_{\infty,1}$.

4.C Connor-Stevens kinetic functions

The ion channel kinetics of the CS models are given by the relationships (4.52), with

$$\alpha_{m,1}(v) = 0.38\frac{-29.7-v}{\exp\left(\frac{-29.7-v}{10}\right)-1} \qquad \beta_{m,1}(v) = 15.2\exp\left(\frac{-54.7-v}{18}\right)$$

$$\alpha_{h,1}(v) = 0.266\exp\left(\frac{-v-48}{20}\right) \qquad \beta_{h,1}(v) = 3.8\frac{1}{\exp\left(\frac{-18-v}{10}\right)+1}$$

$$\alpha_{m,2}(v) = 0.019\frac{-45.7-v}{\exp\left(\frac{-45.7-v}{10}\right)-1} \qquad \beta_{m,2}(v) = 0.2375\exp\left(\frac{-55.7-v}{80}\right)$$

The remaining functions are given by

$$\tau_{m,3}(v) = 0.3632 + \frac{1.158}{\left(1+\exp\left(\frac{v+55.96}{20.12}\right)\right)}$$

$$m_{\infty,3}(v) = \left(0.0761\frac{\exp\left(\frac{v+94.22}{31.84}\right)}{1+\exp\left(\frac{v+1.17}{28.93}\right)}\right)^{\frac{1}{3}}$$

$$\tau_{h,3}(v) = 1.24 + \frac{2.678}{1+\exp\left(\frac{v+50}{16.027}\right)}$$

$$h_{\infty,3}(v) = \frac{1}{\left(1+\exp\left(\frac{v+53.3}{14.54}\right)\right)^4}$$

and

$$\tau_{m,4}(v) = 2.35$$

$$m_{\infty,4}(v) = \frac{1}{1+\exp(-0.15(v+50))}$$

4.D Supporting material

Lemma 4 ([58], Theorem 7.7.7)**.** *Let $H \in \mathbb{C}^{p\times q}$ be a Hermitian matrix given by*

$$H = \begin{bmatrix} A & B \\ B^* & C \end{bmatrix}$$

with $A \in \mathbb{C}^{p\times p}$ and $C \in \mathbb{C}^{q\times q}$. Then H is positive-definite if and only if A is positive-definite and $C - B^*A^{-1}B$ is positive-definite.

Chapter 5

Identification of neuronal systems

In this chapter, we focus on the general problem stated in the beginning of the thesis: identifying nonlinear systems given by the feedback interconnection of a passive system and a system with a fading memory (Figure 2.1). We continue to work with the input additive noise setup (Figure 2.7). Assuming the existence of a finite-dimensional realization of the system in global normal form, and assuming that the data is generated by a forward-Euler discretization of that realization, we thus focus on identifying

$$v_{k+1} = v_k + \tfrac{t_s}{c}\left(-y_k + i_k\right) \tag{5.1}$$

with

$$x_{k+1} = x_k + t_s f(v_k, x_k) \tag{5.2a}$$

$$y_k = h(v_k, x_k) - e_k \tag{5.2b}$$

where $v_k \in V \subseteq \mathbb{R}$, $x_k \in X \subseteq \mathbb{R}^{n_x}$, $y_k \in \mathbb{R}$. We will assume throughout that the sampling time t_s is small enough to ensure that (5.2) realizes an operator with a fading memory. Recall that we refer to (5.2) as the *internal system*, and to (5.2a) as the *internal dynamics*. In this chapter, we use the notation of Sections 4.2 and 1.1.

The purpose of the identification problem is to obtain a simulation model for the system (5.1)-(5.2), that is, to obtain a model that is able to predict the output v given a realization of the input $i_k + e_k$. To do so, we will again rely on the "direct approach" of closed-loop system identification (see Section 2.5.1): we identify the inverse system (5.2) using an output-error model structure.

This chapter is aimed at exploring two extensions of the identification method of Chapter 4. First, we ask whether the output feedback $i_k = \gamma(r_k - v_k)$ used in the last

chapter ("voltage-clamp") can be put aside in favour of exciting the system directly through the applied current i_k ("current-clamp"). Second, we propose an identification method for neuronal systems which does not assume any prior knowledge of the system dynamics.

The feedback requirement

Given the drastically different signals that can be generated by different realizations of the input noise (Figure 2.8), it is important to ensure that the cost function used for identification be asymptotically independent of the realization of the noise. This is crucial to identify the same model regardless of the realization of the input noise. In Chapter 4, we discussed how output feedback can be used to ensure this property: under the assumption that the input additive noise is white, it allowed us to verify the basic assumptions of the PEM method.

It is interesting to consider in which cases we may forgo the output feedback strategy and still verify those assumptions. There are two sources of motivation behind this question. First, from a practical viewpoint, obtaining measurements with voltage-clamp (with output feedback) is more challenging than doing so with current-clamp: it is more difficult to obtain good quality measurements of the voltage while applying feedback at the same time [43].

Second, from the viewpoint of modeling bias, there is the fact that using output feedback in the data collection part of an identification experiment results in an experimental condition that may be rather different from the operating regime of the system when no output feedback is applied. We know from system identification theory that, when the ground truth system does not belong to the model structure, in order to maximize the predictive power of the identified model, it is important that the experimental condition used for data collection be similar to the conditions under which the model will be used for prediction [77, 76].

In Section 5.1, we will show that a noise-induced synchronization phenomenon originally discovered in chaotic systems driven by noisy inputs may allow us to apply the convergence result of the PEM without resorting to output feedback. We will then use this phenomenon to justify an identification method based on a black-box model structure for the internal system (5.2), which we discuss next.

A black-box internal model structure

In Chapter 4, we assumed that the internal dynamics (5.2a) (the ion channel kinetics, in conductance-based and Markov models) was known a priori. Even though this is

a common assumption made by neuroscientists who need to find models to explain neurophysiological data, it is somewhat unrealistic to expect that it holds in a practical situation. It is known, for instance, that there exists a high variability in kinetic parameters that are identified based on neurons of the same type but from different animals of the same species (see [87, 99] and references therein).

To tackle this issue, we completely drop any prior knowledge assumption on the structure of the ground truth internal dynamics. We propose to identify the system in a two-step procedure sharing similarities with the procedure presented in Chapter 3.

In the first step, we use feedback in experiment design, as well as a linear model structure, to identify the eigenvalues of the linearized internal dynamics at the bifurcation equilibrium underlying the system's excitability properties. If the ground truth system has an exponentially contracting internal dynamics, then (based on the previous chapter) we can ensure that this step provides converging estimates of those eigenvalues. In the second step, the identified eigenvalues are used to define a nonlinear black-box model structure possessing universal approximation properties; this model structure is then used to identify the inverse of the neuronal system.

As we have seen in Section 2.5.2, the type of bifurcation and the system dynamics in the proximity of the bifurcation determine much of the excitable behavior of a neuron. We will show that the two-step nature of the procedure is warranted by the fact that using those eigenvalues in the model structure allows for precise identification of the linearized dynamics with relatively few states in the model.

Assuming that the noise-induced synchronization phenomenon occurs for the class of inputs used in the identification experiment, we are able to ensure convergence of the associated cost function. Since we do not assume that the neuronal system belongs to the black-box model structure, we leave the question of asymptotic consistency aside, and focus on model validation. When the system to be identified does not belong to the model structure, then the process of identification becomes one of approximation of the ground truth system. In that case, (asymptotic) bias in the parameter estimates is unavoidable.

Forgoing the output feedback in the nonlinear identification part of the procedure, and working with signals that more closely resemble the behavior of neurons in the brain, allow us to obtain approximations that have good predictive power. We introduce a novel general-purpose signal aimed at identifying the types of excitability discussed in Section 2.5.3, and we demonstrate the identification procedure using artificial data gathered from ground truth models representative of different types of excitable behavior.

The remainder of the chapter is organized in the following way: in Section 5.2, we introduce the model structure used for identifying the neuronal system's inverse model. In Section 5.3, we develop our identification method, including a linear identification stage, a nonlinear identification stage, and input design. In Section 5.4, we demonstrate that the identification method is capable of delivering neuronal models that can quantitatively and qualitatively predict the behavior and excitability type of Class I, Class II, Class II* and bursting neurons.

5.1 Noise-induced synchronization

When two identical nonlinear oscillators are driven by the same periodic excitation, it may happen that their outputs synchronize: they oscillate with the same waveform and phase. An analogous phenomenon may happen when two identical nonlinear systems are driven by the same realization of filtered white noise. In this section, we first introduce Lyapunov exponents, and connect them to the concept of exponential contraction used in the previous chapters. We then review evidence that has been gathered in the physics community showing that the maximal Lyapunov exponent of a driven system can acquire negative values due to noisy excitation. We finish the section by linking the phenomenon to the identification of neuronal systems, using it to justify the application of the PEM to non-equilibrium closed-loop systems.

5.1.1 Lyapunov exponents and synchronization

Consider the system

$$x_{k+1} = f(x_k) + u_k \tag{5.3}$$

where $x_k \in X \subseteq \mathbb{R}^{n_x}$, f is a continuously differentiable function, and u_k is a stochastic forcing term.

In the deterministic case, where $u = 0$, Lyapunov exponents provide a measure of the sensitivity of a trajectory x_k to changes in its initial condition x_0, and a positive maximal Lyapunov exponent gives an indication of chaotic behavior [111, 140]. When u is a nonzero noisy forcing term, the Lyapunov exponents are rather interpreted as a local measure of synchronization between trajectories affected by the same realization of the noise [111]. We will briefly discuss this latter interpretation.

Consider two infinitesimally close trajectories of the system, x_k and $x_k + \delta x_k$, for $k \geq 0$. If δx_k remains sufficiently small, then, to first-order approximation, its evolution

is governed by

$$\delta x_{k+1} = \tfrac{\partial f}{\partial x}(x_k)\delta x_k \tag{5.4}$$

Note that since the trajectory $x_k = \phi_{k,0}(u, x_0)$ depends on x_0 and u (through (5.3)), the linearized dynamics above also depends on those variables.

Given an initial δx_0, the trajectory of (5.4) is given by

$$\delta x_k = M_k(x_0; u)\delta x_0 \tag{5.5}$$

where

$$M_k(x_0; u) = \tfrac{\partial f}{\partial x}(x_{k-1})\tfrac{\partial f}{\partial x}(x_{k-2})\ldots\tfrac{\partial f}{\partial x}(x_0) \tag{5.6}$$

A Lyapunov exponent at x_0 in the direction δx_0 for a trajectory of (5.4) is defined by (see e.g. [21, Definition 2.100] and [6, Definition 3.2.2])

$$\lambda(x_0, \delta x_0; u) = \limsup_{k\to\infty} \frac{1}{k} \log \frac{\|M_k(x_0; u)\delta x_0\|}{\|\delta x_0\|} \tag{5.7}$$

Because the Lyapunov exponents above depend on the noise realization u, in the physics literature they are sometimes called *conditional* Lyapunov exponents [108, 109].

Assuming that $\partial f/\partial x(x_k)$ is bounded for all $k \geq 0$, then $\lambda(x_0, \delta x_0; u) \in \mathbb{R} \cup \{-\infty\}$ for all $\delta x_0 \in \mathbb{R}^{n_x}$, and the number of distinct values that $\lambda(x_0, \delta x_0; u)$ can take as a function of $\delta x_0 \neq 0$ is at most equal to n_x, the dimension of the state-space [6, pp. 144-145]. The largest of those values is called the *maximal Lyapunov exponent*, which we denote $\lambda_1(x_0; u)$. Thus

$$\lambda_1(x_0; u) \geq \lambda(x_0, \delta x_0; u) \tag{5.8}$$

for all δx_0.

In case there exists a $\beta > 0$ such that $\lambda_1(x_0; u) < -\beta < 0$, then it is known that the system (5.4) is exponentially stable [6]. In other words, there exists a $c > 0$ such that

$$\|\delta x_k\| \leq c\,\|\delta x_0\| e^{-\beta k}, \qquad k \geq 0 \tag{5.9}$$

which shows that the two infinitesimally close trajectories initially considered converge to each other — they synchronize due to the common input.

The constant $c > 0$ in (5.9) depends on the realization of the input u, but it also depends in general on the initial condition x_0. The behavior of $\lambda(x_0, \delta x_0; 0)$ as a function of x_0 and δx_0 is the subject of ergodic theory [34, 111]. In particular, there exist results which provide conditions for $\lambda(x_0, \delta x_0; 0)$ to be almost everywhere (with

respect to some measure) independent of x_0, and also for (5.7) to hold with the limit superior replaced by a simple limit.

Rather than delve into this subject, which is out of the scope of this work, we will work with the assumption that synchronization occurs with a constant $c = c(u)$ that is independent of initial conditions. In this case, 5.9 becomes exponential contraction, with the caveat that it holds only for a fixed realization of the input u. This implies global convergence of trajectories subject to the same input realization:

Lemma 5. *Consider the prolonged system* (5.3)-(5.4) *for a fixed realization of* u*, and assume that there exists a* $\beta > 0$ *and a* $c > 0$*, dependent on* u *but independent of* x_0*, such that* (5.9) *holds for all* $x_0 \in X$*. Suppose also that the state-space* X *is convex. Let* $x_k = \phi_{k,0}(u, x_0)$ *and* $\tilde{x}_k = \phi_{k,0}(u, \tilde{x}_0)$*. Then*

$$\|x_k - \tilde{x}_k\| \leq c \, \|x_0 - \tilde{x}_0\| e^{-\beta k}, \qquad k \geq 0$$

for all $x_0, \tilde{x}_0 \in X$*.*

Proof. See Appendix 5.A.1. □

5.1.2 The effect of noise on Lyapunov exponents

An important discovery made over the recent decades is that the presence of a sufficiently strong noisy input u in (5.3) can make the sign of the maximal Lyapunov exponent λ_1 become negative. This may happen in situations where the unforced system ($u = 0$) has $\lambda_1 \geq 0$, which tipically happens for autonomous systems with limit cycles or chaotic attractors. This phenomenon was first observed by physicists when studying the effects of noise in chaotic systems [112, 65, 141], see [12] for a review. It was then found that the same phenomenon occurred in various kinds of neuronal models [64, 154, 36].

This phenomenon, named *noise-induced synchronization*, can be observed in a wide range of excitable and chaotic systems. However, as far as we know, a characterization of the class of systems (and inputs) for which the phenomenon always takes place is currently unknown. It has been shown, for instance, that while the Lorenz chaotic system exhibits noise-induced synchronization when forced by a zero-mean white noise of sufficiently high variance, the same type of noise never induces synchronization in the chaotic Rossler system, no matter how large the noise variance is [12].

One explanation for noise-induced synchronization is that noise drives the trajectories of the system to spend more time on regions of contraction, in the state-space, than in regions of expansion [153]. When the phenomenon occurs in noisy neuronal

oscillators, some authors have observed that it can be studied heuristically in terms of phase-response curves [44, 36].

We illustrate the noise-induced synchronization phenomenon using the Hodgkin-Huxley model in Figures 5.1-5.3. In Figure 5.1, two trajectories, which are initialized at different initial conditions, were obtained by exciting the system with a constant current of $i = 10$ μA/cm^2. In this regime, there is a stable limit cycle in the system dynamics. The trajectories are eventually separated by a constant phase shift, which depends on the initial conditions. The fact that trajectories in the limit cycle never converge to each other is explained mathematically by the existence of a zero maximal Lyapunov exponent[1].

In Figure 5.2, we repeat the experiment, starting the trajectories at different initial conditions, but this time adding a white Gaussian noise signal of zero mean and $\sigma = 10$ μA/cm^2 to the same constant $i = 10$ μA/cm^2. It can be seen that the two trajectories eventually converge to each other – they synchronize, due to the common added white noise at the input.

An important question is whether any realization of the white noise results in a negative Lyapunov exponent. For oscillators that can be studied using the phase reduction method, analytical results have been obtained demonstrating that this is indeed the case [141, 65]. For conductance-based models in the excitable regime (at rest for zero excitation), no analytical results are available.

We can nonetheless perform numerical experiments to check whether different noise realizations cause maximal Lyapunov exponents to become negative. In Figure 5.3, we show how the maximal Lyapunov exponent evolves in time[2] for twenty different trajectories of the Hodgkin-Huxley model, obtained by exciting the system with a constant $i = 7$ μA/cm^2 perturbed by twenty different realizations of added white Gaussian noise e of standard deviation $\sigma = 30$ μA/cm^2. The system is in the excitable regime, with noise causing the system to produce spikes at random times. In all cases, the initial condition was set at the equilibrium state corresponding to $i = 7$ μA/cm^2. It can be seen that in all exponent trajectories, the graph becomes negative, giving an indication that noise-induced synchronization may hold for all realizations of the noise.

[1]This is the real part of the marginally stable Floquet exponent [21] of the periodic time-varying system given by the linearization of the Hodgkin-Huxley model around a trajectory in the limit cycle.

[2]To compute the exponents numerically, we used a Matlab implementation of the algorithm described in [151].

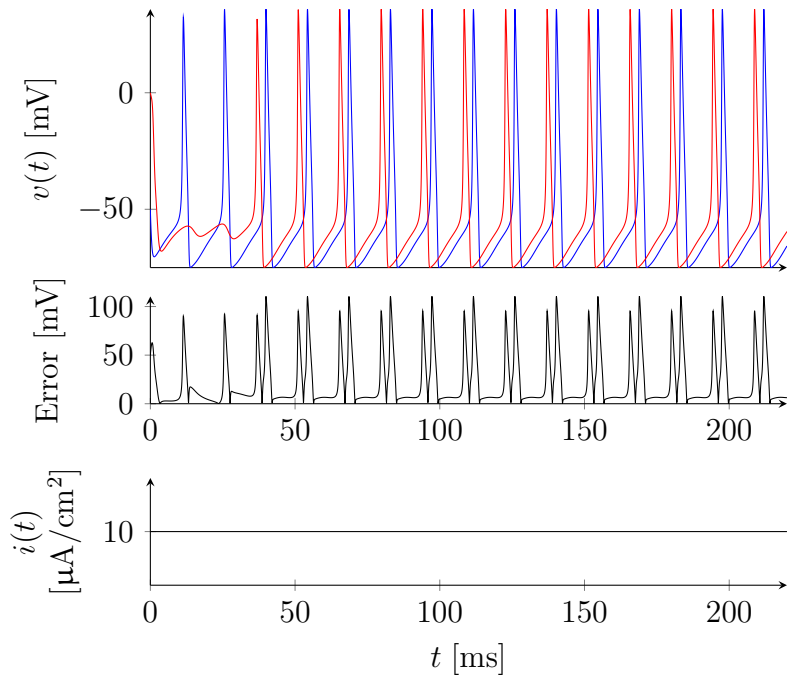


Fig. 5.1 Two different voltage trajectories of the Hodgkin-Huxley model (Figure 2.2) with $i(t) = 10$ μA/cm^2 and no input additive noise.

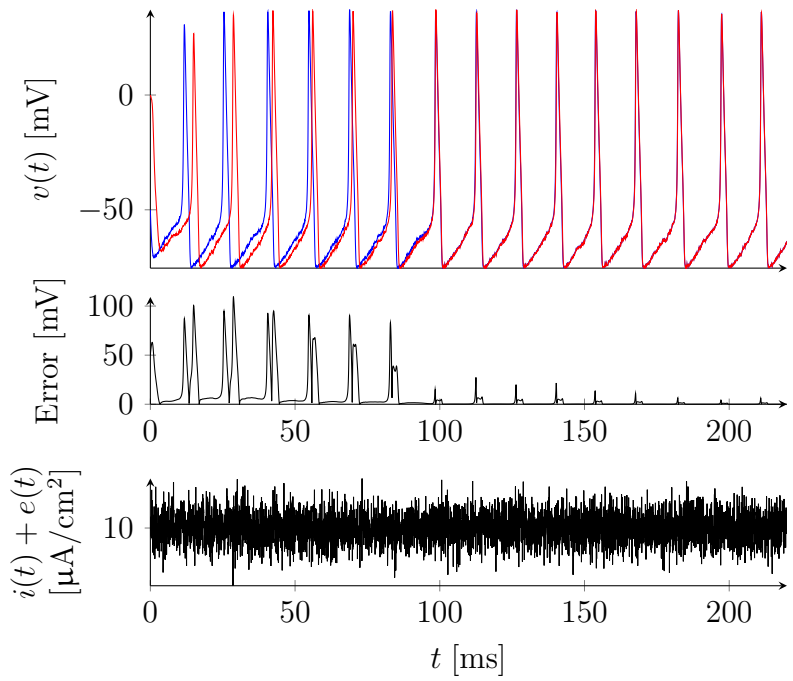


Fig. 5.2 Two different voltage trajectories of the Hodgkin-Huxley model (Figure 2.2) with $i(t) = 10$ μA/cm^2 an input-additive white Gaussian noise $e(t)$ of $\sigma = 10$ μA/cm^2.

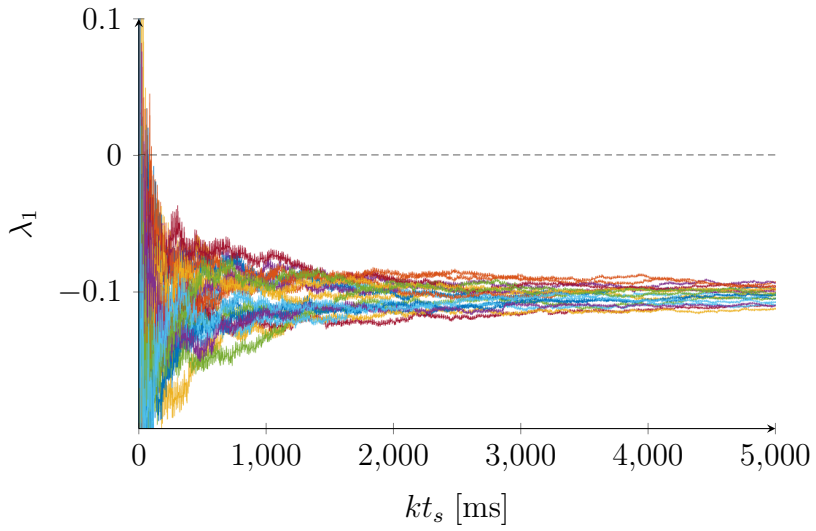


Fig. 5.3 Evolution of the maximal Lyapunov exponent computed numerically for twenty different trajectories of the Hodgkin-Huxley model. The trajectories are obtained by exciting the system with a constant $i = 7$ perturbed by twenty different realizations of input-additive white Gaussian noise e of standard deviation $\sigma = 30$ μA/cm^2.

5.1.3 Noise-induced synchronization and system identification

There are two ways in which noise-induced synchronization can benefit the identification of a neuronal system. First, it may guarantee the fundamental assumptions of the PEM method (Section 2.3.1). Second, it may help identified models to successfully simulate the behavior of the ground truth system subject to noisy data. We develop the first point in this section, and come back to the second point later, when we discuss the validation of identified models.

The next result will be useful to show that the PEM can be applied when global noise-induced synchronization occurs:

Proposition 8. *Consider the system* (5.1)-(5.2)*. Assume that the state-space $V \times X$ is convex, closed and bounded, and that f and h are continuously differentiable. Assume also that i is a deterministic signal in $\mathcal{U}_d$, and that e_k, $k \geq 0$, are zero-mean independent random variables such that all realizations of e belong to $\mathcal{U}_d$.*

Finally, assume that, for all realizations of the noise e, there exists a $\beta > 0$ and a $c > 0$, independent of x_0 (but possibly dependent on i_k and on the probability distributions of e_k) such that the infinitesimal trajectories $(\delta v_k, \delta x_k^\top)$ evolving around the solutions of (5.1)-(5.2) *satisfy*

$$\|(\delta v_k, \delta x_k^\top)\| \leq c \, \|(\delta v_0, \delta x_0^\top)\| e^{-\beta k}, \qquad k \geq 0$$

for all $v_0 \in V$ *and* $x_0 \in X$.

Let

$$\begin{aligned} u_k &= v_k = \frac{t_s}{c}\left(-\sum_{n=0}^{k-1} y_n + r_k\right) + v_0 \\ r_k &= \sum_{n=0}^{k-1} i_n \end{aligned} \tag{5.10}$$

Then Assumption 1 is satisfied for the closed-loop system (5.2), (5.10).

This result can be proven in a completely analogous way to Proposition 6, so we omit its proof. The main difference is that we use Lemma 5, instead of Lemma 2, to show that the past of the inputs is forgotten exponentially fast.

Proposition 8 shows that the main convergence result of the PEM (Lemma 1) may still be used to identify non-equilibrium nonlinear systems, as long as synchronization is induced by the known excitation signal and the input-additive white noise. It may be hard to rigorously prove that a given class of inputs will induce synchronization in a given class of nonlinear systems, in the sense of Lemma 5. Nevertheless, the evidence reviewed in the previous section shows that there are many systems for which it is reasonable to expect noise-induced synchronization to hold.

5.2 Model structure

To identify the internal system (5.2), we will not assume prior knowledge of the structure of the internal dynamics (5.2a). To tackle the problem, we rather rely on the fading memory property of the internal system. It will motivate us to define a black-box model structure based on a *universal approximator*.

The universal approximator we use is given by the cascade interconnection of a finite-dimensional LTI system $G(z)$ and an artificial neural network static nonlinearity $\psi(\cdot)$. Figure 5.4 shows how this model structure is inserted in a feedback loop in order to provide a simulation model for (5.1)-(5.2).

In this section, we define this model structure in detail, and show in what sense it can approximate the system. In Section 5.2.1, we recall how time-delay networks can be used to uniformly approximate a system operator with a fading memory. In Section 5.2.2, we introduce the model structure, which is based on a modification of the time-delay neural network where the delays are replaced by a set of generalized orthogonal basis functions (GOBFs). The choice of the poles of the GOBFs, in view of the closed-loop nature of the identification problem, is discussed in Section 5.3.2.

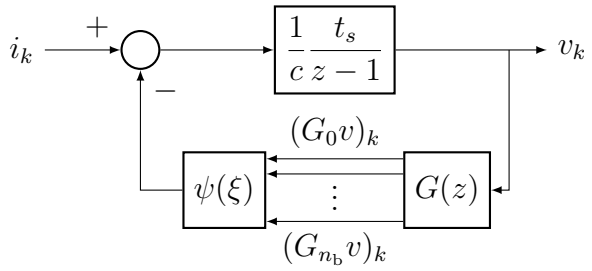


Fig. 5.4 Black-box model structure

5.2.1 Time-delay neural network approximators

System operators that have a fading memory can be uniformly approximated by simple classes of nonlinear systems. A popular class among those is the class of time-delay neural networks (TDNN), which we now describe.

We start by defining an artificial static neural network. We denote by $n_\ell \in \mathbb{N}$ the number of layers in the network. For $\ell = 1, \ldots, n_\ell$, let $n_a^{(\ell)}$ be the number of activation functions in the ℓ^{th} hidden layer, and $n_a^{(0)}$ and $n_a^{(n_\ell+1)}$ the number of inputs and outputs of the network, respectively. Given the weight matrices $W^{(\ell)} \in \mathbb{R}^{n_a^{(\ell+1)} \times n_a^{(\ell)}}$ and the bias vectors $b^\ell \in \mathbb{R}^{n_a^{(\ell+1)}}$, let

$$y^{(\ell)} = W^{(\ell)} u^{(\ell)} + b^{(\ell)}, \qquad \ell = 0, 1, \ldots, n_\ell \tag{5.11}$$

with

$$u_a^{(\ell)} = \frac{1}{1 + \exp\left(-y_a^{(\ell-1)}\right)}, \qquad a = 1, \ldots, n_a^{(\ell)}, \quad \ell = 1, \ldots, n_\ell \tag{5.12}$$

We define an artificial neural network nonlinearity $\psi : u \mapsto y$ by setting $u^{(0)} = u$ and $y = y^{(n_\ell)}$, so that that $n_a^{(0)} = \dim(u)$ and $n_a^{(\ell+1)} = \dim(y)$.

The following result shows that a discrete-time system operator with a fading memory can be approximated uniformly arbitrarily well by a single-layer time-delay neural network.

Lemma 6. *Consider a causal, time-invariant operator $F : \ell_\infty(\mathbb{Z}_+) \to \mathbb{R}(\mathbb{Z}_+)$ with a fading memory on $\mathcal{U}_d$. For an integer $n > 0$, let $H^{(n)} : \ell_\infty(\mathbb{Z}_+) \to \mathbb{R}(\mathbb{Z}_+)^{(n+1)}$ be defined by*

$$(H^{(n)} v)_k = [v_k, v_{k-1}, \ldots, v_{k-n}]^\top, \tag{5.13}$$

and let ψ be a neural network with $n_\ell = 1$, $n_a^{(0)} = n + 1$ and $n_a^{(n_\ell+1)} = 1$. Then, there are integers $n, n_a^{(1)} > 0$ and real parameters $W^{(0)}, W^{(1)}, b^{(0)}, b^{(1)}$ such that for any $\epsilon > 0$,

$$|(Fv)_k - \psi((H^{(n)}v)_k)| < \epsilon, \qquad k \in \mathbb{Z}_+$$

for all $v \in \mathcal{U}_d$.

Proof. See [120, Theorem 1] and [107, Proposition 1]. □

Lemma 6 suggests that the model structure of Figure 5.4 could be realized with $G(z)$ replaced by the transfer function

$$H^{(n)}(z) = [1, z^{-1}, \ldots, z^{-n}]^\top \tag{5.14}$$

and with a single-layer static neural network $\psi(\cdot)$. Although theoretically justified, this choice comes with two issues related to the size of the model structure, that is, its number of states and parameters. First, to accurately approximate an operator with a *slowly* fading memory (e.g., one realized by an exponentially contracting dynamics with a small contraction rate α), the system $H^{(n)}(z)$ should have a *long* (but finite) memory, implying that n should be large. A large n results both in a large number of states in the dynamics of the model structure, and in a large number of parameters in $W^{(0)}$. The second issue relates to the single-layer assumption on the neural network: it is known that increasing the number of layers in an artificial neural network may dramatically decrease the number of elements in the network required to achieve a desired approximation accuracy [138].

A popular approach to overcome the memory limitations of feedforward time-delay networks is to introduce recurrency in the network. This comes at a price: recurrent networks are harder to train, may face instability issues, and possess more local minima than feedforward networks.

5.2.2 Universal approximator with Generalized Orthonormal Basis Functions

We could alternatively interpret the set of delays $\{z^{-n}\}_{n\in\mathbb{Z}_+}$, the elements of $H^{(n)}(z)$ in (5.14), as a set of basis functions in a LTI system space. Depending on the system to be approximated, there might exist a more efficient set of basis functions to perform the approximation, that is, a set which requires fewer parameters in the model to achieve the same approximation margin.

To improve on the TDNN model structure, we replace the set of time-delays by a far more general and flexible set of of LTI basis functions: the so-called set of Generalized Orthonormal Basis Functions (GOBFs) [54, 55]. The aim of this section is to show that the dynamic neural network obtained with GOBFs retains the approximating capabilities of TDNNs.

To define the set of GOBFs, we first have to choose a sequence $\xi = \{\xi_0, \xi_1, \xi_2, \dots\}$ of (possibly complex) poles such that $|\xi_i| < 1$ for all $i \in \mathbb{Z}_+$. Using this sequence, we define the causal transfer functions

$$\begin{aligned} G_0(z) &= z^d \frac{\sqrt{1-|\xi_0|^2}}{z-\xi_0} \\ G_i(z) &= z^d \frac{\sqrt{1-|\xi_i|^2}}{z-\xi_i} \prod_{j=0}^{i-1} \frac{1-\bar{\xi}_j z}{z-\xi_j}, \quad i = 1, 2, \dots \end{aligned} \tag{5.15}$$

with $d = 0$ or $d = 1$.

Note that when $d = 1$ and all the poles in $\{\xi_i\}_{i\in\mathbb{Z}_+}$ are equal to zero, we recover the set of time-delay basis functions. When $d = 0$ and all the poles in $\{\xi_i\}_{i\in\mathbb{Z}_+}$ are the same, we obtain a set of basis transfer functions called Laguerre filters. The utility of generalized basis functions comes from the fact that poles are allowed to be distinct, and thus the choice of poles may be adapted to the system which is being approximated.

The set of GOBFs get their name from the fact that, under a mild condition on the sequence of poles $\{\xi_i\}_{i\in\mathbb{Z}_+}$, they form a fundamental[3] orthonormal set in $\mathcal{H}_2(\mathbb{C}_+)$, the space of functions which are analytic on $\mathbb{C}_+ = \{z \,:\, |z| > 1\}$ and square integrable on the unit circle [96]. In other words,

$$\sum_{k=0}^{\infty} (g_i^*)_k (g_j)_k = 0, \quad \forall\, i \neq j, \quad i, j > 0$$

where here the asterisk denotes complex conjugation, and $(g_i)_k$ are GOBF impulse responses at time k. This property is relevant in the context of neuronal system identification. Figure 5.5 illustrates that for GOBFs defined with timescales which are much larger than those of a hypothetical input spike with voltage trace v_s, the responses $g_i * v_s$ to the spike are similar to the impulse responses g_i. This means that, for relatively slow GOBFs, a fast input spike behaves approximately as an impulse, and thus the outputs of different slow GOBFs to the same fast spike will be close to orthogonal. In the model structure which we will define, orthogonality of the signals at

[3] A fundamental set in a normed space X is a subset $M \subset X$ whose span is dense in X [69].

the output of the GOBFs ensures that the ANN nonlinearity receives a rich (*persistently exciting*) set of inputs to nonlinearly combine.

The set of GOBFs may also form a basis for other system spaces, as the following result shows:

Lemma 7 ([4])**.** *The set* (5.15)*, with* $d = 1$*, is fundamental in* ℓ_1*, the space of discrete-time LTI systems whose impulse responses are absolutely integrable, if* $k^{-\alpha} = O(1 - |\xi_k|)$ *for some* $0 < \alpha < 1$*.*

The next result generalizes Lemma 6 so that GOBFs can be used to approximate a system with a fading memory.

Theorem 3. *Consider a causal, time-invariant operator* $F : \ell_\infty(\mathbb{Z}_+) \to \mathbb{R}(\mathbb{Z}_+)$ *with a fading memory on* $\mathcal{U}_d$*. Let* $\{\xi_i\}_{i\in\mathbb{Z}_+}$ *be a sequence of poles satisfying the condition stated in Lemma 7. For an integer* $m > 0$*, let* $G^{(m)} : \ell_\infty(\mathbb{Z}_+) \to \mathbb{R}(\mathbb{Z}_+)^{(m+1)}$ *be defined by*

$$(G^{(m)}v)_k = [(G_0v)_k, (G_1v)_k, \ldots, (G_mv)_k]^\top \tag{5.16}$$

with the G_i *being the operators associated to* (5.15) *with* $d = 1$*. Let* ψ *be a neural network with* $n_\ell \geq 1$*,* $n_a^{(0)} = m+1$ *and* $n_a^{(n_\ell+1)} = 1$*. Then, there are integers* $m, n_a^{(1)}, \cdots, n_a^{(n_\ell)} > 0$ *and real parameters* $W^{(\ell)}$ *and* $b^{(\ell)}$*, for* $\ell = 0, 1, \ldots, n_\ell$*, such that for any* $\epsilon > 0$*,*

$$|(Fv)_k - \psi((G^{(m)}v)_k)| < \epsilon, \qquad k \in \mathbb{Z}_+ \tag{5.17}$$

for all $v \in \mathcal{U}_d$*.*

Proof. Let $\epsilon_0 > 0$ be given. By Lemma 6, there is a number $n = n(\epsilon_0) > 0$, and parameters $n_a^{(1)} > 0$ and $W_{\text{fir}}^{(0)} \in \mathbb{R}^{n_a^{(1)} \times n+1}$ defining a single-layer artificial neural network ψ such that

$$|(Fv)_k - \psi((H^{(n)}v)_k)| < \epsilon_0, \qquad k \in \mathbb{Z}_+$$

Now let $W_{\text{obf}}^{(0)} \in \mathbb{R}^{n_a^{(1)} \times m+1}$, and consider the vector of transfer functions

$$Q^{(n,m)}(z) = W_{\text{fir}}^{(0)} H^{(n)}(z) - W_{\text{obf}}^{(0)} G^{(m)}(z)$$

By Lemma 7, for every n and every ϵ_1, there exists an $m = m(n, \epsilon_1) > 0$, and a certain $W_{\text{obf}}^{(0)}$ such that

$$\|q_i^{(n,m)}\|_{\ell_1} = \sum_{k=0}^{\infty} |(q_i^{(n,m)})_k| \leq \epsilon_1, \qquad i = 1, \ldots, n_a^{(1)},$$

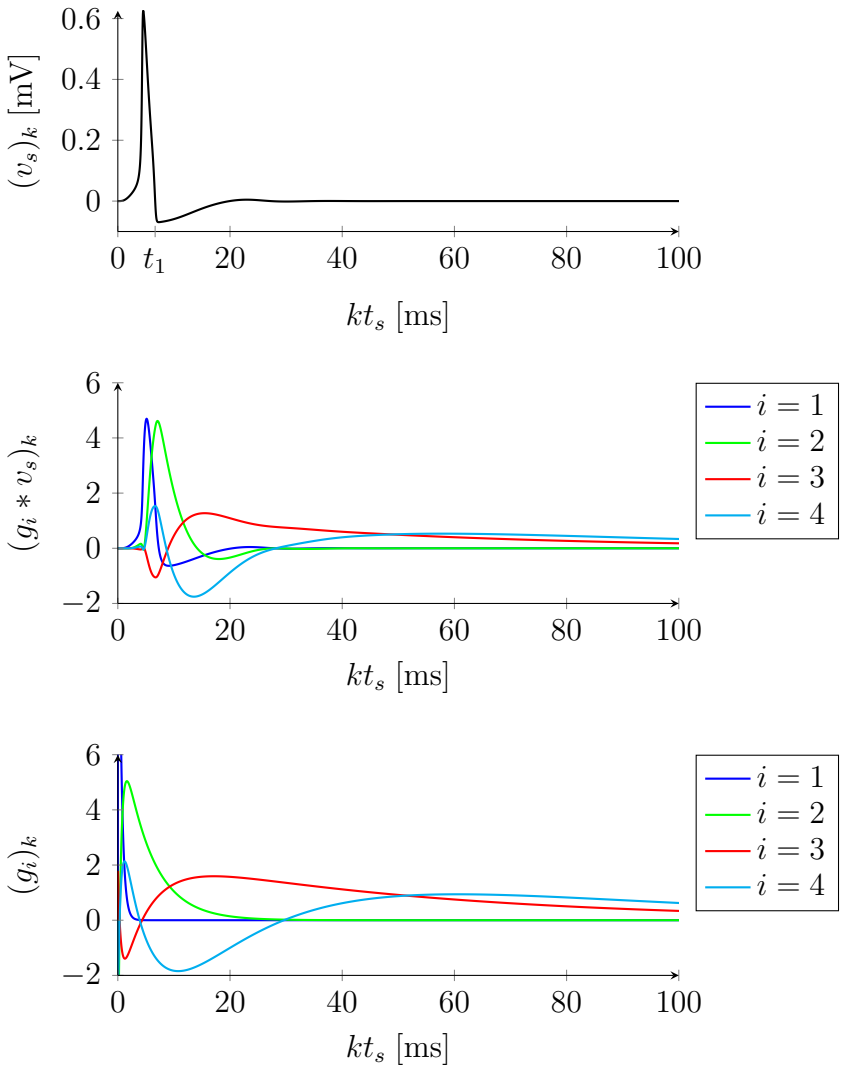


Fig. 5.5 Comparison between impulse responses (bottom) and spike responses (middle) of four GOBFs given by (5.15) with $\xi_1 = 0.9802$, $\xi_2 = 0.9980$, $\xi_3 = 0.9998$ and $\xi_4 = 0.9995$. The spike (top) used to compute the responses $g_i * v$ was normalized so that the area between $t = 0$ and $t = t_1$ equals one (the same area of a Dirac impulse). The impulse responses g_i are mutually orthogonal, while the responses $g_i * v_s$ are not. However, the responses $g_3 * v_s$ and $g_4 * v_s$ are close to g_3 and g_4, respectively, and thus are close to being orthogonal to each other. This can be explained by the fact that the timescales of g_3 and g_4 are much larger than those contained in the spike. The sampling time used is $t_s = 0.01$ ms.

where $q_i^{(n,m)}$ is the impulse response of the i^{th} element of $Q^{(n,m)}(z)$.

Since the activation functions (5.12) are globally Lipschitz, there exists a constant l_0 such that,

$$\begin{aligned}|\psi((H^{(n)}v)_k) - \psi((G^{(m)}v)_k)| &\leq l_0 \|W_{\text{fir}}^{(0)}(H^{(n)}v)_k - W_{\text{obf}}^{(0)}(G^{(m)}v)_k\|_\infty \\ &= l_0 \|(Q^{(n,m)}v)_k\|_\infty \\ &\leq l_0 \max_i \|q_i^{(n,m)}\|_{\ell_1} \|v\|_{\ell_\infty}, \\ &\leq l_0\, \epsilon_1 \beta\end{aligned}$$

for all $k \in \mathbb{Z}_+$. Here, we have used the fact that, for an arbitrary LTI system operator $Q : \ell_\infty \to \ell_\infty^n$, $\max_{i=1,\ldots,n} \|q_i\|_{\ell_1}$ is the induced system gain [155], and that the elements of $\mathcal{U}_d$ are uniformly bounded by β.

Now we have that,

$$\begin{aligned}|(Fv)_k - \psi((G^{(m)}v)_k)| &\leq |(Fv)_k - \psi((H^{(n)}v)_k)| + |\psi((H^{(n)}v)_k) - \psi((G^{(m)}v)_k)| \\ &\leq \epsilon_0 + l_0\, \epsilon_1 \beta\end{aligned}$$

for all $k \in \mathbb{Z}_+$. Given $\epsilon > 0$, we can always find n and m large enough so that we can set $\epsilon_0 < \epsilon/2$ and $\epsilon_1 < \epsilon/(2l_0\beta)$, and obtain (5.17).

This proves the result for a single-layer neural network ($n_\ell = 1$), which was obtained by modifying the input weights $W^{(0)}$ of the time-delay neural network of Lemma 6. The fact that the result also holds for a multi-layer network follows directly from the fact that multi-layer networks with continuous activation functions are capable of arbitrarily accurate approximation (in the uniform norm) to any continuous function over a compact set [59, Theorem 2.1]. A multi-layer network can thus be used to approximate the single-layer network ψ above to arbitrary precision, concluding the proof.

□

We note that this result provides a proof of the uniform approximation capabilities of the model structure suggested by [130] to identify systems with a fading memory.

5.2.3 Internal system model structure

Results such as Theorem 3 do not tell us how large the number of parameters defining the size of the approximating structure (m, n_ℓ, etc.) need to be in order to achieve a certain approximating margin ϵ in (5.17). They also do not tell us how to find the

optimal neural network weight and bias parameters that achieve that margin. However, Theorem 3 provides a solid justification for selecting a model structure to identify (5.2):

Assumption 13 (Model structure)**.** Given a finite sequence $\{\lambda_1, \lambda_2, \ldots, \lambda_{n_\lambda}\}$, and an integer $n_{\text{rep}} \geq 1$, let $\xi_0 = 0$ and

$$\xi_{i+jn_\lambda} = \lambda_i, \qquad i = 1, 2, \ldots, n_\lambda, \quad j = 0, \ldots, n_{\text{rep}} - 1 \tag{5.18}$$

define a sequence of poles. Let

$$G(z) = [G_0(z), G_1(z), \ldots, G_{n_\lambda n_{\text{rep}}}(z)]^\top$$

where the SISO transfer functions $G_i(z)$ are given by (5.15), with $d = 1$. Let $\psi(\,\cdot\,;\theta)$ be an ANN static nonlinearity specified by n_ℓ and n_a, with the vector of parameters θ encompassing the elements of the matrices $W^{(\ell)}$ and the elements of the bias vectors b^ℓ. The model structure used to identify the internal system (5.2) is given by

$$\hat{y}_k(\theta) = \psi(G(q)v_k; \theta) \tag{5.19}$$

where q is the forward shift operator, and the vector θ collects all the neural network weights and biases.

The model structure above uses a periodic sequence of poles for the GOBFs. This automatically satisfies the conditions of Lemma 7. This model structure is the element on the feedback path in Figure 5.4, where $n_{\text{b}} = n_\lambda n_{\text{rep}}$. Note that the choice of ξ_0 implies $G_0(z) = 1$; this choice is made merely to emphasize the direct term in the definition of G. The internal system model structure is commonly called a parallel wiener system [128].

5.3 Identification procedure

The identification procedure has three stages: a preliminary stage, in which an equilibrium bifurcation in the system dynamics is detected; a linear system identification stage, where we identify the poles of a linear approximation of the internal dynamics near the detected bifurcation; and a nonlinear system identification stage, where we estimate the parameters of the nonlinear model structure (5.19). The principle behind the two last stages is similar to that of Chapter 3: first, we use feedback to obtain relevant information about the system; then, we use that information to successfully

identify it. We will discuss these three stages in Sections 5.3.1-5.3.3, and finish with a discussion on input design (for the nonlinear stage) in Section 5.3.4.

5.3.1 Preliminary stage: Bifurcation detection

Equilibrium bifurcations can be detected experimentally, as the current at which the system equilibrium loses stability can be reliably measured. A simple procedure to detect such a bifurcation starts by forcing the system trajectories to a stable equilibrium state by applying a certain (usually negative) current $i_{\min}$ for sufficiently long. We can then obtain a rough estimate of the bifurcation current by applying a perturbed ramp current of the form

$$i_k = i_{\min} + \frac{i_{\max} - i_{\min}}{K t_s} k t_s + \eta_k, \qquad 0 \leq k \leq K, \tag{5.20}$$

where K is the length of the ramp in number of samples, t_s is the sampling period, $i_{\max}$ is the maximum current attained by the ramp, and η_k is a white Gaussian noise of zero mean and finite variance known to the user.

The perturbation η_k in (5.20) is meant to drive the trajectory away from a bifurcation as soon as it happens. The perturbation is necessary to reliably detect Hopf bifurcations, where, without the perturbation, the trajectory can remain in the vicinity of an unstable equilibrium well after the bifurcation has occurred, causing the impression that the equilibrium is still stable. Similarly, the perturbation is necessary to reliably detect a saddle-node bifurcation, where the trajectory can remain in the ghost[4] of an annihilated equilibrium well after the bifurcation has occurred, causing the impression that an equilibrium still exists.

Notice that, in practice, the effect of the user-determined perturbation η_k will be the same as the effect of the input-additive noise e_k in (5.2b). Thus, in practice, we can rely on input-additive noise to help us detect the bifurcation, and η_k can be of low or even zero intensity. Regardless of how the perturbation around the deterministic ramp is generated, the time at which the system's behavior noticeably changes can be used to approximately determine the current i^* and voltage v^* at which a bifurcation occurs. This procedure is best illustrated through an example:

[4]The ghost of a saddle-node equilibrium is the region of the state-space where, following a saddle-node bifurcation, the vector field has a very small magnitude – the annihilated equilibrium's influence lingers on, hence the name. In such regions, the trajectories move very slowly (see, for instance, [63, Section 3.3.5] or [140, Section 4.3]).

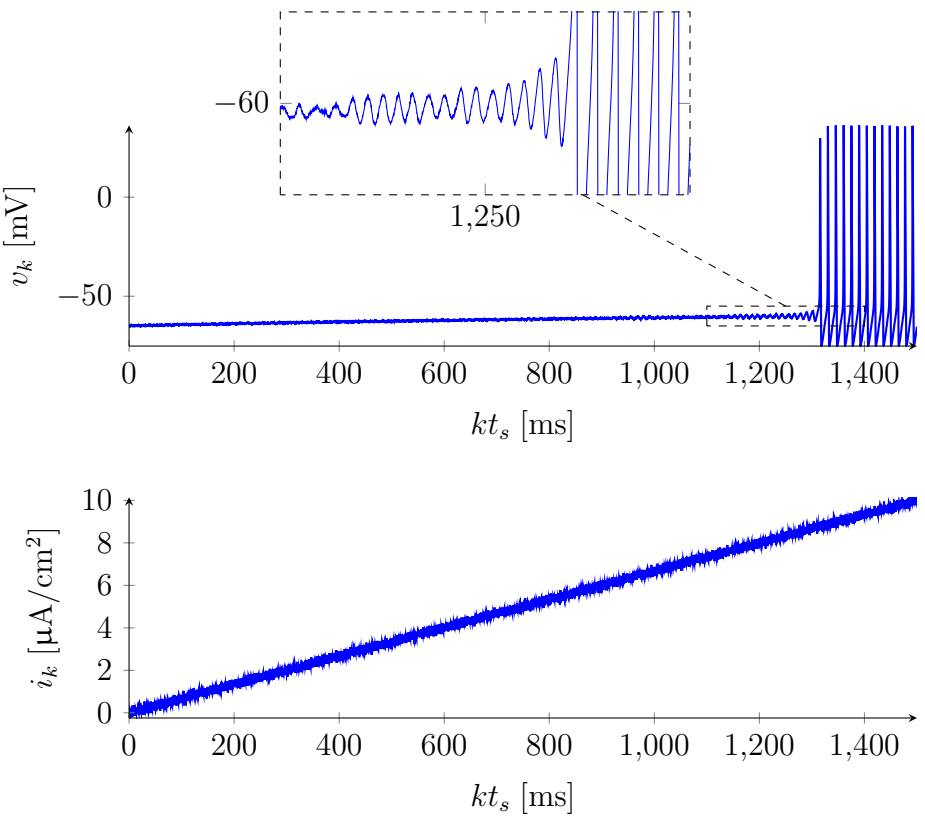


Fig. 5.6 Bifurcation detection experiment for the Hodgkin-Huxley system (data generated using the model described in Example 3 of Chapter 4). Harmonic-type oscillations of increasing amplitude indicate that a Hopf bifurcation has occurred.

Example 8. Figure 5.6 depicts a bifurcation detection experiment in which the Hodgkin-Huxley neuron is excited with a deterministic ramp perturbed by a user-defined white noise of variance 0.1 μA/cm^2 and an input-additive white noise e_k of variance 1 μA/cm^2 (the data was obtained by means of simulations of the HH model described in Example 3 of Chapter 4). It can be seen that around $k^*t_s \approx 1250$ ms, the membrane voltage starts displaying harmonic-type oscillations of increasing amplitude centered around $v^* = -60$ mV; this hints that there is a Hopf bifurcation at that voltage. By the time the neuron starts spiking, the bifurcation has already happened. The bifurcation current can be estimated to be around $i^* \approx 8.3$ μA/cm^2 (this is close to the theoretical value of $i^* = 8.18$ μA/cm^2).

If necessary, to refine the estimate of the bifurcation current, the same procedure can be repeated multiple times, each time using a shallower ramp (this can be achieved by increasing K while keeping $i_{\max}$ and $i_{\min}$ fixed). Alternatively, constant values of i close to the rough value of bifurcation current obtained previously can be used to

probe the stability of the system, in which case the perturbation $\epsilon(t)$ does not need to be used (provided that a long time is allowed for instability to be detected).

5.3.2 Linear identification stage

The purpose of the linear identification stage is to find a good sequence of poles $\{\xi_i\}_{i\in\mathbb{Z}_+}$ for the set of basis functions (5.15). In this section, we use the notation from Section 4.6.1, in which $G_{\text{int}}(z;v)$ denotes the linearization of the internal system (5.2) around the equilibrium point $x_\infty(v)$.

In the discussion that follows, we make use of the following fact:

Proposition 9. *Assume that $G_{int}(z;v^*)$ has a partial fraction decomposition of the form*

$$G_{int}(z;v^*) = a_0^* + \sum_{i=1}^{n_x} \frac{a_i^*}{z-\lambda_i^*}$$

and consider the sequence of GOBFs of Assumption (13)*. Then there is a coefficient vector $C^* = [c_1^*, \ldots, c_{n_\lambda n_{rep}+1}^*]$, such that*

$$|G_{int}(e^{j\omega};v^*) - [a_0^* \quad C^*]G(e^{j\omega})| \leq \sum_{i=1}^{n_x} \left|\frac{a_i^*}{e^{j\omega}-\lambda_i^*}\right| \prod_{m=1}^{n_\lambda n_{rep}+1} \left|\frac{\lambda_i^*-\xi_m}{1-\bar{\xi}_m\lambda_i^*}\right| \tag{5.21}$$

for $\omega \in [0, \pi/t_s]$.

Proof. The coefficients in C^* are obtained by projecting the strictly proper term of $G_{\text{int}}(z;v^*)$ onto the subspace of $\mathcal{H}_2(\mathbb{C}_+)$ spanned by $\{G_i(z)\}_{i=1,2,\ldots,n_\lambda n_{\text{rep}}+1}$. This implies the bound on the right-hand side of (5.21) (see, for instance, [55, Section 2.2]). □

The choice of GOBFs poles

It is natural to first think of an optimal choice of GOBF poles for approximating the ground truth system in an ideal situation where the system dynamics is known. When the ground truth internal dynamics (5.2a) is linear, for instance, the poles of $G_{\text{int}}(z;v) = G_{\text{int}}(z)$ are fixed. In view of Proposition 9, the obvious choice is to use those fixed poles to determine the sequence $\{\lambda_i\}$ of Assumption 13, and set $n_{\text{rep}} = 1$. Those poles would need to be identified, of course, and we will shortly discuss that point.

In general, however, $G_{\text{int}}(z;v)$ will have moving poles and zeros (they will depend on v). In that case, it is not obvious how to choose $\{\lambda_i\}$. We propose that in this situation, a parsimonious choice for the sequence of GOBF poles should be based on

the linearization of the ground truth system at the equilibrium bifurcation determining the type of excitability of the system (see Section 2.5.3).

To make this idea more precise, suppose that, for an input current i^*, there is an equilibrium bifurcation[5] in the dynamics (5.1)-(5.2) at $(v^*, x_\infty(v^*))$. Let $\{\lambda_i^*\}_{i=1,\dots,n_x}$ be the poles of $G_{\text{int}}(z; v^*)$, and let

$$\hat{G}_{\text{int}}(z; v; \theta) = \left.\frac{\partial \psi(u;\theta)}{\partial u}\right|_{u=G(1)v} G(z), \tag{5.22}$$

be the linearization of the model structure (5.19) around the equilibrium $G(1)v$ (note that $G(1)$ is the DC gain of the vector of GOBFs). Then, if we choose $\{\lambda_i\} = \{\lambda_i^*\}_{i=1,\dots,n_x}$, it follows from Proposition 9 that the linearized model structure $\hat{G}_{\text{int}}(z; v^*; \theta)$ will contain the ground truth $G_{\text{int}}(z; v^*)$ even if only one repetition of the sequence of poles in (5.18) is used.

Identifying the GOBF poles

The poles $\{\lambda_i^*\}_{i=1,\dots,n_x}$ of $G_{\text{int}}(z; v^*)$ are initially unknown, and have to be estimated. To do so, we rely on the results presented in Section 4.6: we use output feedback in order to estimate a linear approximation of the nonlinear internal system at the bifurcation point $x^* = x_\infty(v^*)$.

Following the development of Section 4.6, a general strategy to estimate the poles of $G_{\text{int}}(z; v^*)$ is to use a reference input $r_k = \bar{r} + \tilde{r}_k$, $k \geq 0$, such that $v^* = v_\infty(\bar{r})$ and such that $\tilde{r}_k$ is a quasi-stationary signal of small amplitude. Theorem 2 shows that, provided that the relevant assumptions are satisfied, we can obtain an estimate $\hat{G}_{\text{lin}}(z; \hat{\theta}_N)$ of $G_{\text{int}}(z; v^*)$ that converges as $N \to \infty$.

We should not expect $\hat{G}_{\text{lin}}(z; \hat{\theta}_N)$ to be an asymptotically consistent estimate of $G_{\text{int}}(z, v^*)$, since the system behavior is (mildly) nonlinear, and the model structure is linear. However, assuming that the order of $\hat{G}_{\text{lin}}$ is correctly chosen, we should expect some of the poles of $\hat{G}_{\text{lin}}(z; \hat{\theta}_N)$, for N large enough, to approach the corresponding poles of $G_{\text{int}}(z, v^*)$.

For some N, assume that $\{\hat{\lambda}_i\}_{i=1,\dots,\hat{n}_\lambda}$ are the poles of the estimated $\hat{G}_{\text{lin}}(z; \hat{\theta}_N)$ (we omit the subscript N in the $\hat{\lambda}_i$ for clarity). Then, we set $\{\lambda_i\} = \{\hat{\lambda}_i\}_{i=1,\dots,\hat{n}_\lambda}$ so as to define the sequence of GOBF poles (5.18). Note that, as seen in Example 7, in

[5]Note that here we talk about equilibrium bifurcations in discrete-time systems. These bifurcations are different from the bifurcations that occur in continuous-time systems, which were briefly mentioned in Section 2.5.2. For small sampling times t_s, this distinction is of no consequence for the rest of the discussion.

practice, some of the poles of $G_{\text{int}}(z, v^*)$ might be hard to identify (due to zero-pole near cancellations in the ground truth system, or other factors), in which case they are left out of the sequence of GOBF poles.

To correct for the mismatch between $\{\hat{\lambda}_i\}$ and $\{\lambda_i^*\}$, including the possible omission of ground truth poles, we can increase the number of repetitions n_{rep}. From (5.21), this has the effect of decreasing the achievable error in the approximation of $G_{\text{int}}(e^{j\omega}; v^*)$.

We close this section by remarking that this method of choosing basis functions for a model structure containing a linear dynamics can be applied to other types of neuronal modeling paradigms. The GOBFs obtained in this way could be used, for instance, instead of the time-delay basis functions of the Generalized Integrate-and-Fire model described in Section 2.4.2.

5.3.3 Nonlinear Identification stage

Once the poles of the model structure (5.19) have been selected, we can then estimate its parameters. Because we do not assume that the ground truth system belongs to the model structure, we need to identify the system with signals that are representative of its operating regime (such signals will be discussed in the next section). To use such signals, we refrain from applying the feedback method of Chapter 4. Instead, we show that in case noise-induced synchronization occurs during the identification experiment, we are still able to obtain convergence (ergodicity) of the prediction error.

We again make use of the following assumption:

Assumption 14. The signals i_k and $v_k = u_k$ are known (v_k is measured without noise), and i_k is a deterministic signal that belongs to $\mathcal{U}_d$.

Assumption 14 allows us to compute y_k from (5.1).

Theorem 4. *Let $\hat{y}_k(\theta)$ be given by* (5.19) *and y_k be given by* (5.2)*. Let*

$$V_N(\theta) = \frac{1}{N} \sum_{k=1}^{N} \left(y_k - \hat{y}_k(\theta)\right)^2$$

Under Assumptions 13-14, and under the assumptions stated in Proposition 8,

$$V_N(\theta) \to E[V_N(\theta)]$$

uniformly in $\mathcal{D}$, w.p. 1 as $N \to \infty$.

Proof. The minimal state-space realization of the model structure (5.19) has an exponentially contracting dynamics, since the GOBFs are LTI and stable. Since the model structure is only parameterized at the output static nonlinearity, the dynamic part of its state-space realization is the same as that of the partial derivatives of the model structure with respect to the elements of θ. Thus, Assumption 2 is satisfied by the model structure. We can then apply Proposition 8 to show that the ground truth system satisfies Assumption 1. The result then follows from Lemma 1. □

The optimal parameter estimates are obtained as usual through

$$\hat{\theta}_N = \arg\min_{\theta \in \mathcal{D}} V_N(\theta)$$

Notice that, once the GOBFs are fixed, this reduces to the problem of fitting a multi-layer artificial neural network nonlinearity. This can be solved numerically, very efficiently, using the backpropagation algorithm [11]. Although the numerical solution will depend on how parameters are initialized in the algorithm, good solutions can usually be found by solving the problem multiple times, using different initial parameters at each time.

As mentioned earlier, however, we cannot expect to obtain consistent estimates. Instead, we now focus on input design and validation.

5.3.4 Input design and validation

When identifying a system that does not belong to the model structure, the input used to perform parameter estimation heavily influences the quality of the identified model; the parameter estimates $\hat{\theta}_N$ are only optimal with respect to the deterministic input i_k used to excite the system.

Even for simple linear systems, it is well known that a model that is estimated and successfully validated on one class of signals may not be well validated on another class of signals (for a simple analytical example, see [77, Example 8.2]). A good type of training input is thus one that results in a model that generalizes well towards (that is validated sucessfully on) different classes of input.

Inputs used in the literature

In the context of biophysically realistic neuronal models, [33] studied how well models trained on either steps, ramps, or constant-mean white noise generalized towards input types not used for training. It was found that models trained on combined data

from multiple experiments with ramp and step inputs generalize well towards noisy input; on the other hand, models trained solely on noisy input do not generalize well towards ramp and step inputs. The conclusion seems to be that using noisy input in the identification of neuronal systems might not be the best choice for input design, and indeed many works in the field of parameter extraction for biophysically realistic neuronal models limit the type of the excitation to either steps or ramps.

However, there is evidence in the field of system identification that filtered white noise (and random phase multisines) are good signals to identify nonlinear oscillators [26]. Since periodic oscillatory behavior is a possible regime of excitable behavior, this evidence should at least in part apply to those systems as well. Furthermore, filtered white noise has an advantage as an input signal over ramps and steps in that it allows arbitrarily long data-gathering experiments to be conducted, whereas when using ramps and steps, multiple experiments of a limited length need to be conducted.

It is worth mentioning that some authors have used swept-sine signals (also known as *chirp* or *zap*), given by sinusoids with increasing frequencies in time, to characterize the frequency response of single neurons [118, 110]. One issue with this type of signals, as shown in [26], is the fact that they may not provide a sufficiently rich excitation for nonlinear systems possessing limit-cycles.

Input design for identifying excitability

One class of inputs that does not seem to have been explored in depth in the literature is the combination of user-defined white (or filtered) noise with a deterministic ramp (note that the user-defined input noise in i_k should not be confused with the unmeasured membrane current noise e_k). We have previously described this type of input in Equation (5.20), in the context of bifurcation detection. We reproduce it here for convenience:

$$i_k = i_{\min} + \frac{i_{\max} - i_{\min}}{Kt_s} kt_s + \eta_k, \qquad 0 \leq k \leq K,$$

We argue that the class of inputs above ideal to identify excitability, particularly when using a very general model structure such as the one introduced earlier. The key lies in a judicious selection of the value of the variance of the noise η_k.

Keeping the variance of η_k sufficiently low is useful to ensure that the linearized behavior of the system in the subthreshold voltage range $v < v^*$ is well captured by the estimated model (in Figure (5.6), for instance, this is the range where $v < -60$ mV). By estimating well the linearized behavior of the ground truth system, the estimated model will not have spurious bifurcations of the stable equilibrium that defines the

subthreshold regime of the system. Furthermore, estimating well the linearized behavior guarantees that the right bifurcation is learned by the estimated model.

In the nonlinear identification stage, the variance of η_k should, however, not be too small. First, it is important to keep the signal-to-noise ratio as large as possible. Equally important, the noise should be large enough to allow identification of the limit cycle attractor underlying spiking behavior. To identify such attractors, it is important that the the state-space of the original system be explored as much as possible close to the limit cycle [26], and for that to happen, the input noise level should be sufficiently high.

Periodizing the ramp

In order to design effective input signals to identify excitable behaviors, we need to be able to capture the hysteresis of the F-I curve that defines Type II and Type II* excitability. For that reason, it is important that we not only apply an increasing ramp, but also a decreasing ramp that starts from a high current.

We thus propose an input signal in which white or filtered white noise is added to a deterministic *sawtooth* wave — a periodic combination of increasing and decreasing ramp signals. This input is given by

$$i_k = \begin{cases} i_{\min} + 2\dfrac{i_{\max} - i_{\min}}{K}(k - pK) + \eta_k, & pK \leq k < pK + K/2 \\ i_{\max} - 2\dfrac{i_{\max} - i_{\min}}{K}(k - pK - K/2) + \eta_k, & pK + K/2 \leq k < (p+1)K \end{cases} \tag{5.23}$$

for $p = 1, \ldots, P$, where K is the period of the wave, in number of samples, assumed to be such that $K/2$ is an integer, and P is the number of periods of the wave. The signal η_k is white or coloured noise.

In addition to allowing for the identification of excitability, a periodized ramp allows for arbitrarily long experiments (in which case the convergence results presented before can be applied).

Validation

Despite the fact that our identification strategy focuses on identifying the internal system (5.2) with an output error nonlinear model structure, the ultimate objective of the identification procedure is to be able to successfully *simulate* the excitable behavior of the closed-loop system (5.1)-(5.2), and not just that of the internal system. This means that, for validating the identified model, we need to compare the output of

(5.1)-(5.2) to the output of

$$\begin{aligned} \hat{v}_{k+1} &= \hat{v}_k + \tfrac{t_s}{c}(-\hat{y}(\theta) + i_k + e_k) \\ \hat{y}_k(\theta) &= \psi(G(q)\hat{v}_k; \theta) \end{aligned} \tag{5.24}$$

with $\theta = \hat{\theta}_N$, using the same deterministic input i_k and *the same realization of the noise* e_k, in case noise is used in the validation stage.

As previously discussed in Section 2.5, the behavior of two closed-loop systems containing open-loop elements with quantitatively similar input-output behaviors is not guaranteed to be qualitatively or quantitatively similar unless the open-loop components share the same localized properties. By designing the identification procedure so as to capture the local bifurcation underlying excitability, we expect to successfully validate the closed-loop simulation model (5.24).

At the same time, validation of the closed-loop model can be enhanced by the noise-induced synchronization phenomenon described in Section 5.1. It has previously been observed that noise-induced synchronization can happen if the systems driven by the same noise realization are not identical; this has been referred to as a "structural stability" property of noise-induced synchronization [64, 65] This means that the identified model (5.24) and the ground truth system (5.1)-(5.2) may synchronize when subject to the same realization of noisy input.

5.4 Numerical results

In this section, we apply the identification procedure described in Section 5.3 to various neuronal models that are representative of different excitability classes. Data is generated using models from the literature, corrupted by input-additive noise. The purpose of these numerical examples is to provide a proof of concept of the method, and, at the same time, provide insights into the problem of neuronal system identification.

5.4.1 Repeating GOBF poles smoothly improve IF curve identification

In this section, we identify the Hodgkin-Huxley (HH) model. The continuous-time HH model is detailed in Example 3, and its kinetic functions are given in Appendix 4.B.

To identify the system, we follow the three identification stages described in Section 5.3. When we reach the nonlinear identification stage, we will illustrate the effect of

increasing the number of GOBFs in the model structure of Assumption 13. To do that, we will estimate three models, increasing the number of repetitions in the sequence (5.18) for each model.

We start with the first stage, that of bifurcation detection. In Example 8, we have shown that a single bifurcation detection experiment yields a bifurcation voltage of $v^* \approx -60$ mV. We thus use this value in the next stage.

The second stage in the procedure is that of identifying a linear system close to the detected bifurcation, with the goal of finding GOBF poles. In Example 7 of Chapter 4, we used voltage-clamp to identify several linear models approximating the linear behavior of the internal system of the HH model around specific voltage setpoints. The closest setpoint to the bifurcation detected in the first stage is $\bar{v} = -60.08$ mV . We can use the poles of the linear system identified at that setpoint for defining the GOBFs. The discrete-time linear system identified around $\bar{v} = -60.08$ mV has a pole at $\hat{\lambda}_1 = 0.9981$ and another one at $\hat{\lambda}_2 = 0.9664$. For a sampling time of $t_s = 0.01$ ms, these correspond to the system time constants $\tau_1 = 5.395$ and $\tau_2 = 0.2974$, respectively, shown in Figure 4.8 (we ignore the third, fast spurious pole shown in the figure).

The third step in the procedure is that of nonlinear identification. Defining $\{\hat{\lambda}_i\}_{i=1,2} = \{0.9981, 0.9664\}$, and fixing the neural network size parameters $n_a^{(1)} = n_a^{(2)} = 5$, we investigate the effect of choosing $n_{\text{rep}} \in \{1, 2, 4\}$ in the quality of the identified model. Note that for $n_{\text{rep}} = 1$, the identified model will be a *reduced* model: it has three states, one fewer than the HH model.

To generate data for the nonlinear identification step, we simulated a discretized ground truth Hodgkin-Huxley model with $t_s = 0.01$ subject to the input i_k given by (5.23) with the parameters shown in Table 5.1. Before that excitation is applied, the system is allowed to relax to the equilibrium corresponding to i_{min}. The data used in the nonlinear identification stage is shown in Figure 5.7. The input-additive noise used in the experiment is white Gaussian noise such that $\sigma[e_k] = 1$ μA/cm^2 for all $k \geq 0$.

Model	i_{min}	i_{max}	K	$\sigma[\eta_k]$	$\sigma[e_k]$	t_s	P
HH	-20	50	20×10^5	20	1	10^{-2}	2
CS Class I	0	40	20×10^5	20	1	10^{-2}	2
CS Class II*	68	93	20×10^5	20	1	10^{-2}	2

Table 5.1 Parameters of the deterministic input i_k in (5.23) and of the unmeasured noise e_k used for simulating the identification experiments in Sections 5.4.1 and 5.4.2. The terms η_k and e_k are white Gaussian noise with the indicated standard deviation for all $k \geq 0$. Current units are in μA/cm^2; K is the number of samples; t_s is in ms, and P is the number of periods of the input.

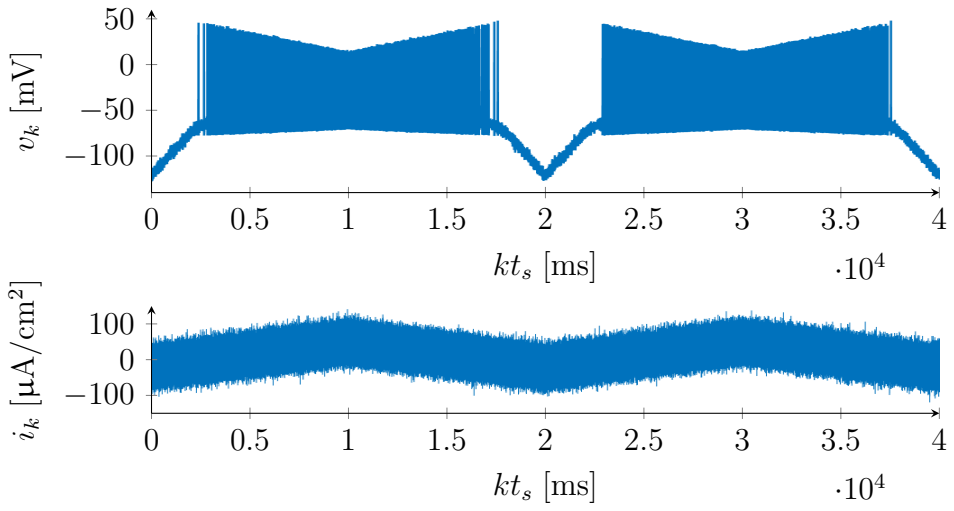


Fig. 5.7 Data used for training in the nonlinear stage of the identification of the Hodgkin-Huxley system, as described in Section 5.4.1. Top: membrane voltage. Bottom: applied current.

Figure 5.8 show that increasing the number of poles in the model structure smoothly improves identification of the FI curve for the same training data. Figures 5.9 and 5.10 show that the models, that were estimated based on the sawtooth wave, generalize well towards coloured noise inputs. To finish this section, we show in Figure 5.11 that the success in the identification of the excitable behavior of the HH model is tied to the success in the identification of its localized behavior close to a bifurcation (see the discussion in Section 2.5.2).

5.4.2 Continuous and hysteretic F-I curves

In this section, we identify a modified version of the Connor-Stevens neuron [23, 31]. This neuron possesses three currents: a fast inward inactivating sodium current, a slow outward potassium current, and a slow outward inactivating A-current. The ground truth continuous-time model is given by

$$\begin{aligned}\dot{x} = &-120m_1^3h_1(v-55) - 20m_2^4(v+75)\\ &-\bar{g}_{\mathrm{A}}m_3^3h_3(v+75) - 0.3(v+17) + i + e\end{aligned} \tag{5.25}$$

with internal dynamics given by Equations (4.9a)-(4.9b) together with the kinetic functions detailed in Appendix 4.C. The constant $\bar{g}_{\mathrm{A}} > 0$ is of particular relevance to this example.

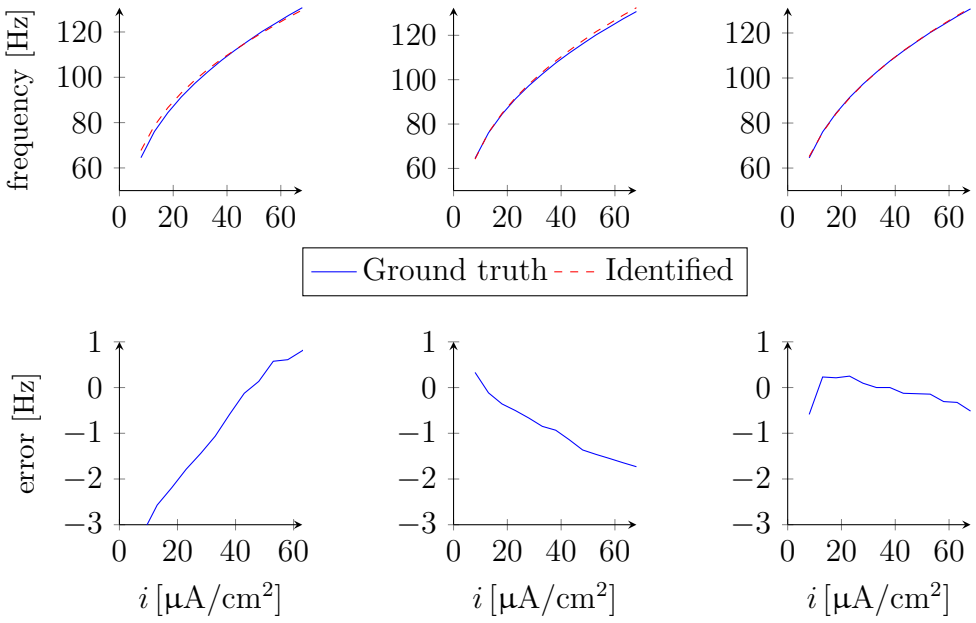


Fig. 5.8 Frequency-response curves of identified Hodgkin-Huxley systems: $n_{\text{rep}} = 1$ (left), $n_{\text{rep}} = 2$ (middle), $n_{\text{rep}} = 4$ (right). In all cases, $n_a^{(1)} = 5$ and $n_a^{(2)} = 5$.

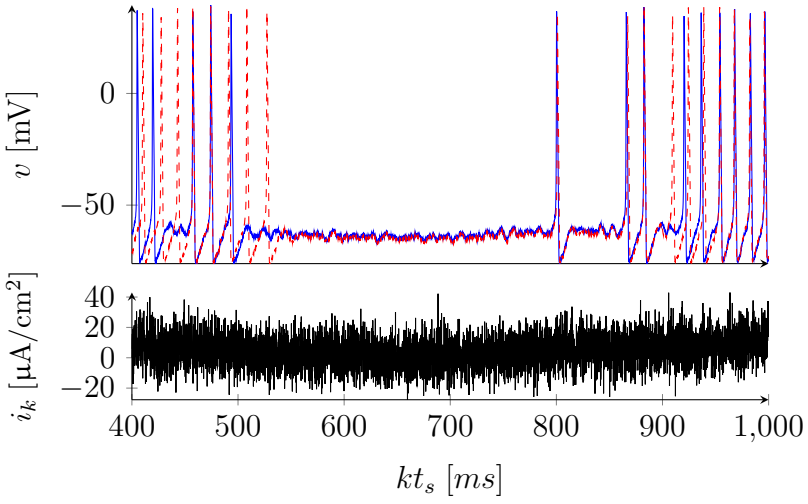


Fig. 5.9 Top: voltage of HH neuron (blue, solid) and voltage of identified model (red, dashed) with $n_{\text{rep}} = 1$, $n_a^{(1)} = 5$, $n_a^{(2)} = 5$.

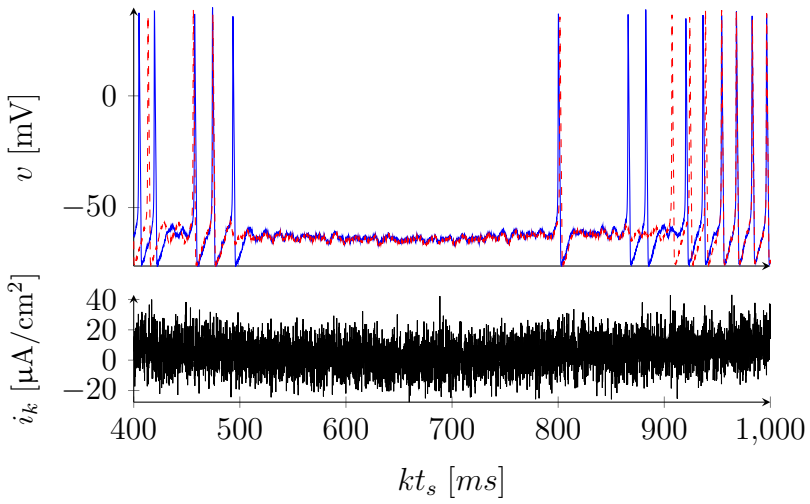


Fig. 5.10 Top: voltage of HH neuron (blue, solid) and voltage of identified model (red, dashed) with $n_{\text{rep}} = 2$, $n_a^{(1)} = 5$, $n_a^{(2)} = 5$.

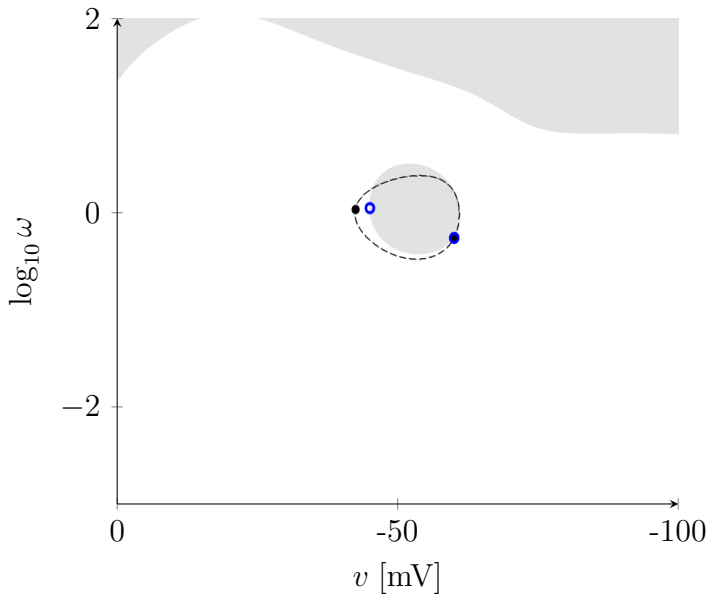


Fig. 5.11 The set ∂G (dashed) of the ground truth continuous-time HH model is superimposed on the regions of local positive conductance (white) and local negative conductance (grey) of the identified model with $n_{\text{rep}} = 2$ and $n_a^{(1)} = n_a^{(2)} = 5$ (see Section 2.5.2 for an introduction to these concepts). Black dots represent Hopf bifurcations in the HH model; blue circles represent bifurcations in the discrete-time identified model. The bifurcation leading to spike initiation ($v^* = -60.2$) is correctly identified. The bifurcation leading to spike termination could be better identified by increasing $i_{\max}$ in the excitation signal. Note that imprecisely identifying the spike-terminating bifurcation does not negatively impact validation of the model for the input current range of interest (Figure 5.10).

The sodium and potassium currents are similar to those found in the Hodgkin-Huxley neuron. In the absence of the A-current ($\bar{g}_{\mathrm{A}} = 0$), the model (5.25) possesses a Type II excitable behavior (based on a subcritical Hopf bifurcation). The A-current can change the excitability type of the modified Connor-Stevens neuron. This is because the inactivating gating variable of the A-current introduces a slow positive feedback component to the system dynamics.

When the maximal conductance of the A-current is $\bar{g}_{\mathrm{A}} = 90$, the model (5.25) bifurcates through a saddle-node on invariant circle bifurcation around $i^* = 25.7$, $v^* = -59.7$, and displays type I excitability. Increasing the A-current maximal conductance to $\bar{g}_{\mathrm{A}} = 250$, the system bifurcates through a saddle-node (off invariant circle) bifurcation around $i^* = 92.9$, $v^* = -60.8$, and displays robust rest-spike bistability. This is a characteristic typical of type II* excitability [30].

For simplicity, in this example we use the ground truth poles of $G_{\mathrm{int}}(z; v^*)$ in order to select the GOBF poles of the model structure. In other words, we set $\{\hat{\lambda}_i\} = \{\lambda_i^*\}$. The continuous time constants associated to those poles are listed in Table 5.2. The discrete-time poles of $G_{\mathrm{int}}(z; -59.7)$ and $G_{\mathrm{int}}(z; -60.8)$ are obtained through the forward-Euler relation $\lambda_i = 1 - t_s/\tau_i$.

To identify the Connor-Stevens model in the class I configuration ($\bar{g}_{\mathrm{A}} = 90$), and in the class II* configuration ($\bar{g}_{\mathrm{A}} = 250$), we simulated the respective discretized models with $t_s = 0.01$, subject to the input i_k given by (5.23) and the input parameters shown in Table 5.1. The data used for training in the nonlinear identification stage is shown in Figures 5.12 and 5.13. The model structure poles were chosen as the poles of $G_{\mathrm{int}}(z; -59.7)$ in the class I case, and the poles of $G_{\mathrm{int}}(z; -60.8)$ in the class II* case; in both cases, we used a single repetition $n_{\mathrm{rep}} = 1$ for the sequence of poles, and a two-layer static neural network with 10 units in each layer ($n_a^{(1)} = n_a^{(2)} = 10$).

Model	v^*	τ_1	τ_2	τ_3	τ_4	τ_5	τ_6	τ_7	τ_8	τ_9	τ_{10}
CS I	-59.7	0.04	1.86	2.96	0.99	2.97	-	-	-	-	-
CS II*	-60.8	0.04	1.79	2.98	1.01	3.01	-	-	-	-	-
STG	-49.7	0.12	1.78	5.57	8.92	26.75	6.39	52.21	20	13.54	220.5

Table 5.2 Continuous-time ground truth time constants of the internal dynamics of various models at the respective bifurcation points v^*. The GOBF poles of the discrete-time model structures are computed according to $\lambda_i = 1 - t_s/\tau_i$, with t_s the sampling time of the simulation. CS I: Connor-Stevens Class I; CS II*: Connor-Stevens Class II*; STG: Stomagastric ganglion neuron.

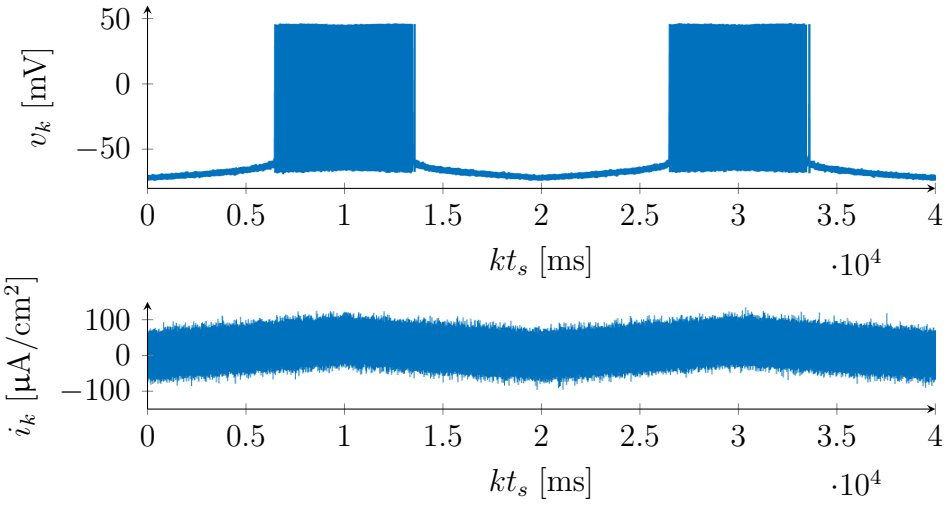


Fig. 5.12 Data used for training in the nonlinear stage of the identification of the Connor-Stevens class I system, as described in Section 5.4.2. Top: membrane voltage. Bottom: applied current.

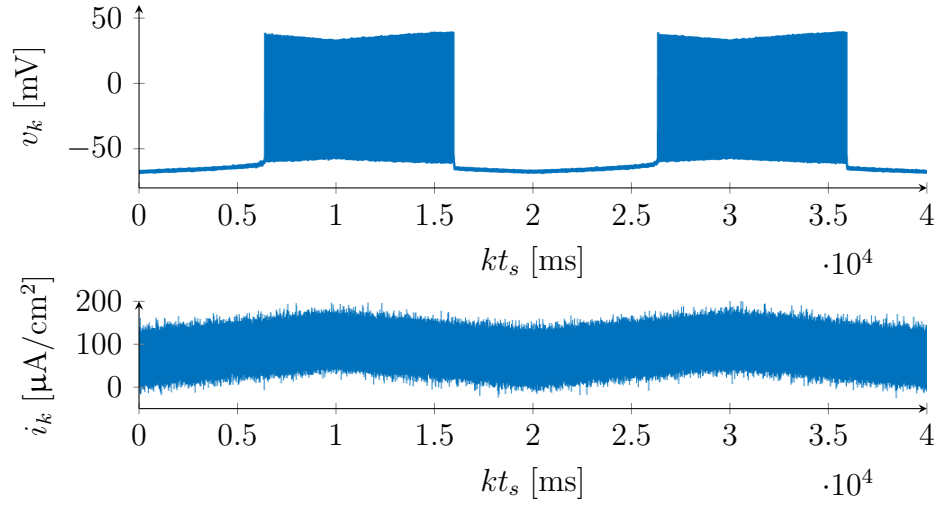


Fig. 5.13 Data used for training in the nonlinear stage of the identification of the Connor-Stevens class II* system, as described in Section 5.4.2. Top: membrane voltage. Bottom: applied current.

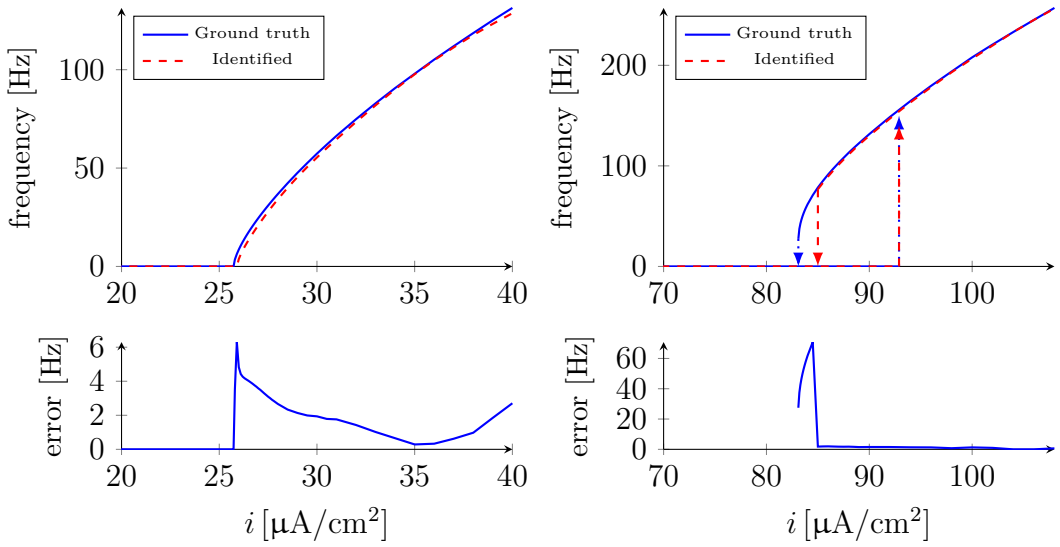


Fig. 5.14 Frequency-response curves of identified Connor-steven neurons: Type I excitable configuration (left) and type II* excitable configuration (right). In both cases, $n_{\text{rep}} = 1$, $n_a^{(1)} = n_a^{(2)} = 10$.

To corrupt the data, we set the input-additive unmeasured noise to white Gaussian noise with $\sigma[e_k] = 1$ μA/cm^2 for all $k \geq 0$. To discard transient data, we removed the first 0.5 seconds of the simulation from the dataset.

Figure 5.14 shows that the identified models are able to qualitatively and quantitatively capture the right class of excitability in both cases. The only noticeable inaccuracy is around the spiking-to-rest part of the hysteric F-I curve of the class II* model. To improve the identification of the $F-I$ curve of the class II* model, we could either use more data in the cost function (by adding additional periods to the input i_k), and/or increase the number of GOBFs and neural network layers in the model structure.

Figures 5.15 and 5.16 show that, with minor inaccuracies, the same identified models are able to successfully reproduce simulated spiking data. The effect of the inaccuracy in the identified F-I curve of the class II* model can be seen around 800 ms in Figure 5.16: the identified model stops spiking earlier than the ground-truth model.

5.4.3 Bursting

In this section, we identify a neuron from the crab stomagastric ganglion (STG), a system responsible for producing rhythmic muscular activity in the crab's stomach [86]. STG neurons are capable of bursting autonomously, a behavior different from those considered in the previous examples. In fact, autonomous bursting can be thought of

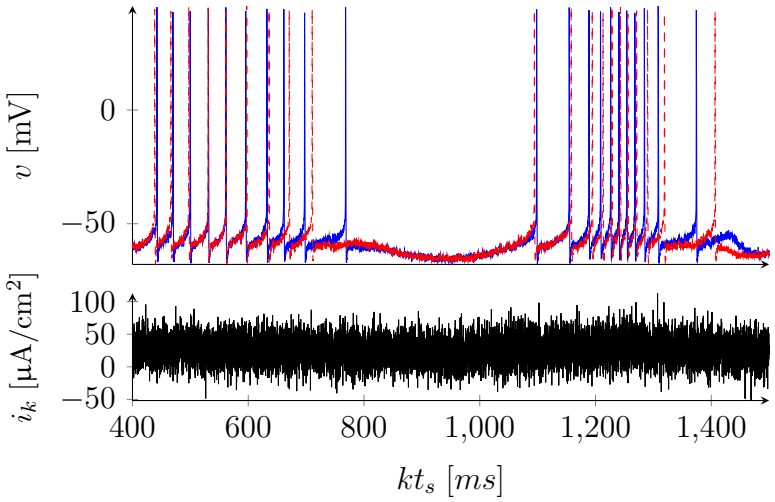


Fig. 5.15 Top: voltage of modified Connor-Stevens neuron in type I configuration (blue, solid) and voltage of identified model (red, dashed) with $n_{\text{rep}} = 1$, $n_a^{(1)} = n_a^{(2)} = 10$ (red, dashed). Bottom: input i_k to both systems.

as class II* excitability which is modulated by an ultra-slow negative feedback current [42].

The STG model used in this example is given in continuous-time by

$$\begin{aligned}\dot{v} =& -700\, g_{\text{Na}}(v-50) - 80\, g_{\text{K}}(v+80) - 30\, g_{\text{A}}(v+80) - 25\, g_{\text{KCa}}(v+80)\\ &- 6\, g_{\text{CaT}}(v-80) - 9\, g_{\text{CaS}}(v-80) - 0.1\,(v+50) + i + e\\ =& -i_{\text{Na}} - i_{\text{K}} - i_{\text{A}} - i_{\text{CaT}} - i_{\text{CaS}} - i_{\text{KCa}} - i_{\text{L}} + i + e\end{aligned} \tag{5.26}$$

where $g_{\text{Na}}, \ldots, g_{\text{CaS}}$ are voltage-dependent adaptive variables.

The STG model has two qualitatively different types of ion currents. The first type includes the regular voltage-gated currents i_{Na}, i_{K}, i_{A}, i_{CaT} and i_{CaS}, and the second type accounts for the calcium-activated potassium current i_{KCa}. The regular voltage-gated currents are modeled using the standard Hodgkin-Huxley formalism, that is, they are given by

$$i_{\text{ion}} = \bar{g}_{\text{ion}}\, m^p h^q (v - \nu_{\text{ion}}) \tag{5.27}$$

where the states m and h evolve according to dynamics of the type

$$\begin{aligned}\tau_m(v)\dot{m} &= -m + m_\infty(v)\\ \tau_h(v)\dot{h} &= -h + h_\infty(v)\end{aligned} \tag{5.28}$$

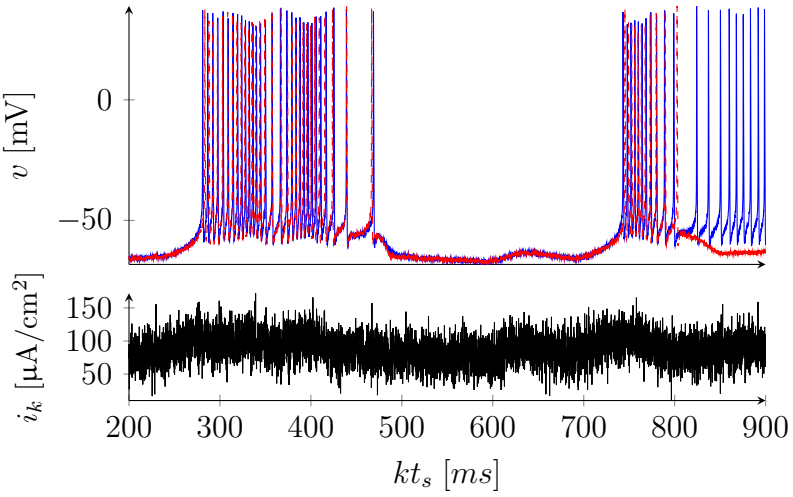


Fig. 5.16 Top: voltage of modified Connor-Stevens neuron in type II* configuration (blue, solid) and voltage of identified model (red, dashed) with $n_{\text{rep}} = 1$, $n_a^{(1)} = 10$, $n_a^{(2)} = 10$ (red, dashed). Bottom: input i_k to both systems.

The dynamics and parameters of the gating variables above are detailed in [75, pp. 2318-2319]. In addition to purely voltage-gated currents, the STG model possesses a potassium current whose conductance is activated by the concentration of calcium ions near the channel. The calcium concentration in turn depends on the intensity of the calcium currents i_{CaS} and i_{CaT}. The i_{KCa} current is given by

$$\begin{aligned} \tau_{\text{KCa}}(v)\dot{m}_{\text{KCa}} &= -m_{\text{KCa}} + m_{\text{KCa},\infty}(v,z) \\ \tau_z\,\dot{z} &= -z - \kappa(i_{\text{CaT}} + i_{\text{CaS}}) + b \\ i_{\text{KCa}} &= 25 m_{\text{KCa}}^4 (v+80) \\ m_{\text{KCa},\infty}(v,z) &= \frac{z}{z+3}\,\frac{1}{1+\exp\left((v+28.3)/-12.6\right)} \end{aligned} \tag{5.29}$$

where $\kappa = 0.94$, $\tau_z = 20$, and $b = 0.05$. The state variable z represents calcium concentration, while m_{KCa} is the gating variable of the calcium-activated current.

In Appendix 5.A.2, we show that the STG neuronal model above has an exponentially contracting internal dynamics on closed and bounded regions of the state-space. We remark that the arguments used in Appendix 5.A.2 can immediately be extended to show that this property holds for any conductance-based model possessing any number of Hodgkin-Huxley type currents (5.27)-(5.28) and calcium-activated currents of the form (5.29). Thus the framework developed in Chapter 4 and in this chapter can also be used to identify the STG neuron, and indeed any bursting neuron with calcium-activated currents.

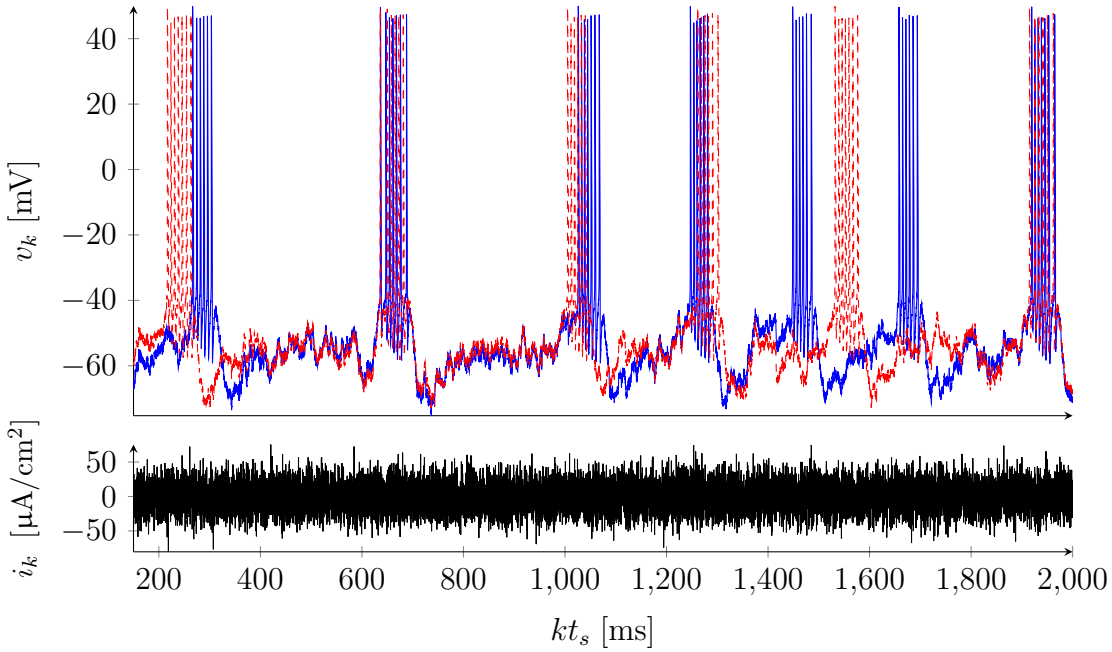


Fig. 5.17 Top: voltage of the ground truth STG model (blue, solid) and voltage of the identified model with $n_{\text{rep}} = 1$, $n_a^{(1)} = 15$, $n_a^{(2)} = 12$ (red, dashed). Bottom: input i_k to both systems.

To construct our model structure, we note that the model above has a bifurcation for $i^* \approx -0.26$, at $v^* = -49.7$. For simplicity, we select the ground truth poles of the linearized internal system $G_{\text{int}}(z; -49.7)$. Those poles are given by $\lambda_i = 1 - t_s/\tau_i$, with the time constants τ_i shown in Table 5.2.

In this example, we focus on identifying the burst excitability of the STG model by subjecting it to an input current given by

$$i_k = -0.5 + \eta_k$$

where η_k is a Gaussian white-noise signal such that $\sigma[\eta_k] = 20$ µA/cm^2. The ground truth model was simulated with $t_s = 0.075$ and subject to the input above for $Nt_s \approx 10$ seconds. To eliminate transient effects, we discard 1 second of data from the dataset.

In Figure 5.17, we can see that the identified model reproduces the bursts of the STG model both quantitatively and qualitatively. To improve these validation results, more data can be added to the experiment, or the model structure order can be increased.

5.5 Conclusion

In this chapter, we introduced a phenomenological black-box model structure which can be tailored to identify the excitable properties of neuronal systems. Using this model structure, we have shown that it is possible to obtain models that qualitatively and quantitatively reproduce the behavior of the ground truth systems being identified.

Because the experiment design proposed in this chapter precludes the use of a stabilizing output feedback, we have justified a rigorous application of the PEM with the phenomenon of noise-induced synchronization.

In our numerical examples, we obtained models possessing a number of states that is comparable to (and sometimes smaller than) that of the ground truth systems. The model structure is flexible, but at the same time it has a very simple parametrization. Because of the feedforward architecture of the model structure, the nonlinear optimization problem resulting from the application of the PEM method is simple to solve numerically.

5.5.1 Relation to existing models

In Section 2.4, we have seen that biophysical models and GLIF models are the two most common types of neuronal models used for estimation. Given the popularity of these models, and the existence of various methods to estimate them, one could question the necessity of introducing yet another model class of neuronal systems.

To answer that, we start by observing that the model that has been introduced in this chapter lies somewhere in between GLIFs and biophysical models in terms of how close to neurophysiology it is. While the open-loop model structure

$$\hat{y}_k(\theta) = \psi(G(q)v_k; \theta)$$

contains a bank of linear filters, a feature that is also present in GLIFs, the model structure is based on the idea that to approximate neuronal behavior, we should primarily approximate the dynamics of ionic currents — the essence of biophysical models. The output $\hat{y}_k$ can in fact be interpreted in neurophysiological terms as the sum of all ionic and leak currents in the neuron.

Something that has not been explored in this chapter is the fact that the model structure above can also be used to approximate *individual* ionic currents. If individual ionic current data is available, it is in principle possible to estimate a model $\hat{y}_k(\theta^{(j)})$ for each ionic current (indexed by j). Adding the outputs of these models together

(following Kichoff's law) yields a model of a neuron's internal dynamics that has an extra layer of biophysical interpretability. It follows that the model used in this chapter can be constrained so as to mimic biophysical models in terms of their architecture.

On the practical side, an advantage of our model over biophysical models is the fact that, starting from a situation where *nothing* is known about a neuron, they can be simpler to estimate. Note that all of the estimation methods mentioned in Section 2.4.1 assume that a parametrization of ion channel kinetics is readily available, and, apart from [60], require solving potentially large and involved optimization problems. The method presented here relies on two simple identification steps (a LTI estimation problem, followed by training an ANN) which can be solved reasonably quick in a personal computer. Furthermore, the linear identification step can in principle be skipped, and GOBF poles can be chosen based on intuition about the timescales of the neuronal behavior.

When compared to GLIFs, an advantage of our model is that it maintains the smoothness that is required for faithfully reproducing the nonlinear dynamics of neurons, including all the relevant bifurcations. For instance, a particular example of nonlinear behavior that has not been observed in GLIFs is a robustly hysteretic F-I curve such as the one observed on the right of Figure 5.14. A robust hysteresis is linked to the existence of slow ionic currents that provide positive feedback to the neuronal membrane, and those in turn have been shown to play a critical role in controlling the participation of individual neurons in the rhythmic behavior of CPGs [28]. Our models are thus good candidates for use in dynamic clamp studies (see Section 2.4.3) where circuit behavior depends critically on the nonlinear dynamics of individual neurons.

In sum, it is fair to conclude that the estimation procedure described in this chapter yields good, smooth models at low computational cost that can replace biophysical models whenever a neuron needs to be identified from zero initial knowledge. Additionally, the model structure above can be used as a simpler-to-train alternative to detailed Markov models of ion channel kinetics (see Section 2.4.1).

5.5.2 Other kinds of inputs

In order to identify the F-I curves of neuronal systems, we argued for exciting the system with an applied current given by the combination of a sawtooth wave and coloured noise. Our work leaves the door open for using other types of excitation, and ultimately the best choice of excitation may depend on the application envisaged for the identified model. In the next chapter, we will also discuss, as future work, the idea of performing optimal input design.

Appendix

5.A Proofs

5.A.1 Proof of Lemma 5

Let

$$\gamma(s) = sx_0 + (1-s)\tilde{x}_0$$

Since $\gamma(s) \in X$, there is a solution $\phi_{k,0}(u, \gamma(s))$ of (5.3) that satisfies

$$\phi_{k+1,0}(u, \gamma(s)) = f(\phi_{k,0}(u, \gamma(s))) + u_k$$

and

$$\frac{\partial}{\partial s}\phi_{k+1,0}(u, \gamma(s)) = \frac{\partial f}{\partial x}(\phi_{k,0}(u, \gamma(s)))\frac{\partial}{\partial s}\phi_{k,0}(u, \gamma(s))$$

The prolonged system above is in the form (5.3)-(5.4). By assumption, there exists a $\beta = \beta(u) > 0$ and a $c = c(u) > 0$ such that

$$\left\| \frac{\partial}{\partial s}\phi_{k,0}(u, \gamma(s)) \right\| \leq c \left\| \frac{\partial}{\partial s}\gamma(s) \right\| e^{-\beta k}, \qquad k \geq 0 \tag{5.30}$$

The flow $\phi_{k,0}(u, \cdot)$ deforms the initial segment $\gamma(s)$, $s \in [0, 1]$ into a curve in state-space. Let L_k be the length of that curve,

$$L_k = \int_{s=0}^{s=1} \left\| \frac{\partial}{\partial s}\phi_{k,0}(u, \gamma(s)) \right\| ds \tag{5.31}$$

We can now see from (5.30) and (5.31) that

$$\|\phi_{k,0}(u, x_0) - \phi_{k,0}(u, \tilde{x}_0)\| \leq L_k \leq c \|x_0 - \tilde{x}_0\| e^{-\beta k}$$

for $k \geq 0$.

5.A.2 Internal dynamics of the STG model

Let x denote the internal states of the system, such that the model is in the form (2.33)-(2.34). Without loss of generality, assume that the states are ordered such that $x = (\ldots, m_{\mathrm{CaT}}, h_{\mathrm{CaT}}, m_{\mathrm{CaS}}, h_{\mathrm{CaS}}, m_{\mathrm{KCa}}, z)^\top$. Then we have

$$\frac{\partial f}{\partial x}(v, x) = \begin{bmatrix} J_{11}(v) & 0 \\ 0 & J_{22}(v, x) \end{bmatrix}$$

with $J_{11}(v)$ a diagonal matrix containing elements of the type $-1/\tau(v)$, and J_{22} given by

$$J_{22} = \begin{bmatrix} -1/\tau_{\mathrm{CaT},m}(v) & 0 & 0 & 0 & 0 \\ \vdots & \ddots & 0 & 0 & 0 \\ 0 & 0 & -1/\tau_{\mathrm{CaS},h}(v) & 0 & 0 \\ 0 & 0 & 0 & -1/\tau_{\mathrm{KCa}}(v) & 1/\tau_{\mathrm{KCa}}(v)\partial m_{\mathrm{KCa},\infty}(v,z)/\partial z \\ -\kappa/\tau_z \partial i_{\mathrm{CaT}}/\partial m_{\mathrm{CaT}} & \cdots & -\kappa/\tau_z \partial i_{\mathrm{CaS}} \partial h_{\mathrm{CaS}} & 0 & -1/\tau_z \end{bmatrix}$$

We can show exponential contraction of the internal dynamics with a block-diagonal matrix $P = \mathrm{diag}(P_{11}, P_{22})$, where $P_{11} = p_0 I$ for some $p_0 > 0$, and P_{22} is the block-diagonal matrix $P_{22} = \mathrm{diag}(p_1 I_4, p_2, p_3) > 0$, where I_4 denotes a 4×4 identity matrix. Computing $J_{22}^\top P_{22} + P_{22} J_{22}$, and applying Lemma 4 (see Appendix 4.D) with

$$A = 2\,\mathrm{diag}(-p_1/\tau_{\mathrm{CaT},m}(v), \ldots, -p_1/\tau_{\mathrm{CaS},h}(v), -p_2/\tau_{\mathrm{KCa}}(v))$$

and $C = -2p_3/\tau_z$, we can see that by choosing $p_3 > 0$ and $p_2 > 0$ sufficiently small we can ensure that $J_{22}^\top P_{22} + P_{22} J_{22}$ is uniformly negative-definite on closed and bounded regions of the state-space X (alternatively, we could choose $p_1 = p_3^2$, and then choose p_3 sufficiently large).

Chapter 6

Conclusion

In this thesis, we took steps towards bridging the gap between a classical system identification method, the prediction error method, and the field of estimation of neuronal models. The prediction error method (and system identification in general) relies on a fundamental fading memory assumption. At the same time, neuronal models and (more generally) non-equilibrium systems lack this property. Our focus has been to analyze in which ways the difficulties posed by the lack of a fading memory can be circumvented in order to effectively identify neuronal models.

6.1 Recapitulation

We now provide some take-home messages related to the developments found in the previous chapters.

Output feedback plays a key role in nonlinear system identification

A key observation in this thesis is that any neuronal model can be endowed with a fading memory (or with its state-space counterpart, a contracting dynamics) by output feedback. Although this property has been at the core of electrophysiology since the invention of the voltage-clamp experiment by Hodgkin and Huxley, we were not aware of previous work making use of it in the identification of nonlinear systems with the prediction error method.

More specifically, we have shown the following:

- If the system's internal dynamics is purely linear and the system has a simple Lure feedback structure, output feedback allows the nonlinearity to be identified

consistently (under mild assumptions) separately from the linear dynamic part (Chapter 3).

- Output feedback guarantees consistency of parameter estimates using the PEM under an input additive white noise assumption (Chapter 4).
- If the internal dynamics is nonlinear, then output feedback allows some of the moving poles (with respect to an output setpoint) of the internal dynamics to be identified. However, in this case consistency is no longer obtained (Chapter 4).

In each of the two cases above, the information gathered with "feedback identification" can be used to make the subsequent nonlinear identification step much more efficient.

In the case of a Lure system, knowledge of the nonlinearity can be used to turn the identification problem into a linear problem by the use of feedback linearization (Chapter 3); this knowledge could also be used to identify the system using the indirect method (see Section 2.5.1).

In the case of conductance-based models, knowledge of the moving poles obtained in the feedback experiment can be used in many ways. For instance, it helps us determine the $\tau(v)$ functions of the classical Hodgkin-Huxley-type models without resorting to experimental procedures such as ion-channel blocking. As we have shown in Chapter 5, it is also appealing to use these previously identified poles in the definition of a universal approximator model structure that facilitates the identification of the bifurcations underlying excitability.

Noise-free voltage measurements allow for system identification with additive input current noise

Unmeasured input current noise is an intrinsic feature of neuronal models and a major issue in nonlinear system identification. Under the assumption of negligible voltage measurement noise, we showed that the identification of a neuronal system can be done using the "direct approach" of classical closed-loop identification theory. When this approach is applied to neuronal models, it is the inverse system that is identified. The noise at the input of the closed-loop neuronal system becomes noise at the output of the inverse system. Furthermore, since the inverse system has a fading memory, the output error model used in the identification procedure satisfies the assumptions of the PEM method.

The assumption of negligible measurement noise is of course idealized. However, it does prove the point that, in face of unknown membrane current noise, a good strategy

may be to invest in accurate measurements so as to apply the techniques described in the previous chapters.

A single long experiment is sufficient for system identification

The existence of unmeasured noise at the input leads to different measurements in different experiments, even when the user-defined input remains the same. A common strategy used in the neuroscience community is to estimate models based on datasets that are averaged over multiple experiments (see, e.g, [92] and [46]).

In this work, we provided two distinct approaches that allow for a rigorous asymptotic analysis with a single input signal. In Chapter 4, we have shown that if feedback is used to ensure a fading memory property, then the estimates are asymptotically independent of the noise realization. In Chapter 5, we have exploited the interesting property that noise can itself induce a fading memory, even in the absence of feedback. This is the so-called noise-induced synchronization phenomenon, which provides an alternative mechanism to guarantee the fundamental assumptions of the PEM method.

The identification of excitability must be localized

When using black-box models for the identification of neuronal ion channels (the inverse system), a question arises about the type of input that should be used in the data-gathering experiment. We have argued in this work for the use of inputs that provide enough data for the estimated model to capture the localized behavior of neuronal models in the vicinity of bifurcation points. Capturing this local behavior is a necessary condition for correctly identifying the excitability class of a neuronal system.

To stress that point, we have shown how knowledge of this localized behavior can be incorporated in the identification procedure. Namely, the poles of the linearized dynamics of the internal model (i.e., of the ion currents) in the vicinity of a bifurcation can be used to define an otherwise parsimonious black-box model structure.

6.2 Future research directions

In the next two sections, we briefly discuss two directions for future developments that are suggested by the research developed in this thesis.

Optimal experiment design

In Section 2.4.2, we have seen that, because Generalized Linear Models (GLMs) are defined directly in terms of a likelihood function $\mathrm{p}(s_{[0:k]} \mid u_{[0:k]}, \theta, \vartheta)$, it is possible to efficiently compute the posterior of model parameters on-line, i.e., while gathering data for identification. As a consequence, the input used for exciting the system during the experiment can be adapted on-line so as to maximize the information about the parameters that is contained in the next output sample.

In Section 5.3.4, when dealing with black-box models, we have opted for using an offline approach to input design. While this approach did result in satisfactory estimation results, as shown in the examples at the end of Chapter 5, it does comes with the downside that many of the parameters of the sawtooth input wave (5.23) have to be tuned manually before each identification experiment. This means that more than one experiment might be necessary before an input that results in a good model can be found; furthermore, to select the input parameters successfully, some minimal knowledge about the timescales and excitability regimes of the neuron is required.

It would of course be desirable to improve on this subptimal input design strategy. Given the success of optimal input design algorithms for GLMs, a relevant question is whether that framework can be adapted for biophysical (conductance-based or black-box) neuronal models.

The main obstacle faced when trying to design optimal inputs for biophysical models is the fact that it is not trivial to compute the likelihood $\mathrm{p}(v_{[0:k]} \mid i_{[0:k]}, \theta)$ of a closed-loop model structure (e.g., the model structure (5.24)). First, because a closed-form expression is not available, and second, because it is not immediately obvious whether a numerical method used to compute it would converge.

One possibility that can be explored in the future is to approach the problem from the point of view of identification of stochastic state-space models [124]. For instance, using balanced realizations of GOBFs[1], the model structure (5.24), could alternatively be represented in terms of a state-space system. It may then be possible to apply existing sequential Monte Carlo methods [124, 123] for estimating the system state, as well as the probability distributions which are necessary for computing optimal inputs in real time. The feasability of such an approach, in practice, would also depend on whether posterior updating and mutual information computation could be done within a sampling interval of the identification experiment.

[1] See [55] for details on how such realizations can be constructed.

Identification of detailed neuronal networks

There are deep connections between the feedback strategy used in Chapters 3 and 4 and the theory of passivity and differential passivity. The assumption that the internal dynamics is contracting, uniformly in v, implies that the system is strongly minimum-phase: the zero dynamics, obtained by forcing $v = 0$ in the internal dynamics, has a globally exponentially stable equilibrium point [18]. It is known that a strongly-minimum phase system with a relative degree of one can always be transformed into a strictly passive system by state feedback [18, 134]. A uniformly exponentially contracting internal dynamics thus represents an even stronger form of the minimum-phase property: one which allows making the system strictly passive by output feedback, which does not require measurement of all states. In fact, the output feedback transforms the system into a strictly differentially passive system [39]. The existence of an interconnection theory for such systems leads to interesting questions regarding the identification of coupled neuronal systems.

We have focused most of our work on identifying single compartment neuronal models. However, as we have remarked earlier, neuronal networks can be accommodated in our framework by assuming multivariable signals. An interesting research direction concerns the use of passivity and differential passivity theory to make further developments in the identification of such networks. For instance, passivity and dissipativity theory have been used successfully in the past to study the problem of synchronization of oscillators [116, 132, 139]. As we have seen in Chapter 5, synchronization helps us ensure ergodicity of the prediction error, required to obtain convergent and consistent parameter estimates. Thus, the interconnection theory of passivity theory may provide us with the right angle to tackle the question of how to use the PEM in order to identify networks of coupled neuronal systems.

Noise assumption

In this thesis, we mostly worked with an input-additive (membrane current) noise setup, where the noise was assumed to be white. The noise at the output (the membrane voltage) was ignored, an assumption which allowed us to estimate the internal current (from the derivative of v) and prove consistency of parameter estimates.

Possible research directions involve situations in which these noise assumptions are not satisfied:

- If there is measurement noise in the membrane voltage, then we have to deal with the fact that the regressors used in the least squares problem of Section

4.4.2 become noisy. The problem becomes an Errors-in-Variables identification problem [79]. These problems can be addressed, for instance, using total least squares methods (see [88] for a review).

- Measurement noise in the membrane voltage is amplified by the differentiation operation, used to reconstruct the internal current signal. The indirect closed-loop identification approach (see Section 2.5.1) would offer an alternative that does not require differentiation of the measured voltage. In that case, an output feedback can be applied to the model structure to make its dynamics exponentially contracting (thus ensuring a predictor with an exponentially fast fading memory, as in Assumption 2). Two complications arise from adopting the indirect approach: first, the unknown membrane current noise $e(t)$ has to be treated as process noise (which we avoided with the direct closed-loop approach); second, the optimization problem becomes drastically nonlinear in the model structure parameters.

- If the input-additive noise is coloured (e.g., if it is filtered Gaussian white noise), then the consistency result obtained in Chapter 4 does not hold anymore. One possible extension of the method thus involves the assumption of input-additive (membrane current) coloured noise.

6.3 Final remarks

This thesis establishes a bridge between two classical fields of research: system identification and electrophysiology. We hope that its results will foster further interest from the system identification community to contribute to important questions of neuroscience.

References

[1] Abarbanel, H., Creveling, D., Farsian, R., and Kostuk, M. (2009). Dynamical State and Parameter Estimation. *SIAM Journal on Applied Dynamical Systems*, 8(4):1341–1381.

[2] Abarbanel, H. D. I., Creveling, D. R., and Jeanne, J. M. (2008). Estimation of parameters in nonlinear systems using balanced synchronization. *Physical Review E*, 77(1).

[3] Abbott, L. F. (1999). Lapicque's introduction of the integrate-and-fire model neuron (1907). *Brain Research Bulletin*, 50(5):303–304.

[4] Akçay, H. and Ninness, B. (1998). Rational Basis Functions for Robust Identification from Frequency and Time-Domain Measurements. *Automatica*, 34(9):1101–1117.

[5] Almog, M. and Korngreen, A. (2016). Is realistic neuronal modeling realistic? *Journal of Neurophysiology*, 116(5):2180–2209.

[6] Arnold, L. (1998). *Random Dynamical Systems*. Springer Monographs in Mathematics. Springer-Verlag, Berlin Heidelberg.

[7] Aström, K. J. and Murray, R. M. (2008). *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press, Princeton.

[8] Aylward, E. M., Parrilo, P. A., and Slotine, J.-J. E. (2008). Stability and robustness analysis of nonlinear systems via contraction metrics and SOS programming. *Automatica*, 44(8):2163–2170.

[9] Barclay, A., Gill, P. E., and Ben Rosen, J. (1998). SQP Methods and their Application to Numerical Optimal Control. In Schmidt, W. H., Heier, K., Bittner, L., and Bulirsch, R., editors, *Variational Calculus, Optimal Control and Applications*, pages 207–222. Birkhäuser Basel, Basel.

[10] Billings, S. A. (2013). *Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains*. John Wiley & Sons, Hoboken, New Jersey.

[11] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer-Verlag, New York.

[12] Boccaletti, S., Kurths, J., Osipov, G., Valladares, D. L., and Zhou, C. S. (2002). The synchronization of chaotic systems. *Physics Reports*, 366(1):1–101.

[13] Boyd, S. and Chua, L. (1985). Fading memory and the problem of approximating nonlinear operators with Volterra series. *IEEE Transactions on Circuits and Systems*, 32(11):1150–1161.

[14] Brette, R. (2015). What Is the Most Realistic Single-Compartment Model of Spike Initiation? *PLOS Computational Biology*, 11(4):e1004114.

[15] Burghi, T. B., Schoukens, M., and Sepulchre, R. (2019). Feedback for nonlinear system identification. In *2019 18th European Control Conference (ECC)*, pages 1344–1349.

[16] Burghi, T. B., Schoukens, M., and Sepulchre, R. (2020a). Feedback Identification of conductance-based models. *Automatica*. In press; available at: https://arxiv.org/abs/2002.09626.

[17] Burghi, T. B., Schoukens, M., and Sepulchre, R. (2020b). System identification of biophysical neuronal models. In *59th IEEE Conference on Decision and Control*, Jeju Island, Republic of Korea. To appear.

[18] Byrnes, C. I., Isidori, A., and Willems, J. C. (1991). Passivity, feedback equivalence, and the global stabilization of minimum phase nonlinear systems. *IEEE Transactions on Automatic Control*, 36(11):1228–1240.

[19] Cannon, R. C. and D'Alessandro, G. (2006). The Ion Channel Inverse Problem: Neuroinformatics Meets Biophysics. *PLOS Computational Biology*, 2(8):e91.

[20] Casas, R. A., Bitmead, R. R., Jacobson, C. A., and Johnson, C. R. (2002). Prediction error methods for limit cycle data. *Automatica*, 38(10):1753–1760.

[21] Chicone, C. (2006). *Ordinary Differential Equations with Applications*. Texts in Applied Mathematics. Springer-Verlag, New York, 2nd edition.

[22] Ciraula, M. and Sandberg, I. W. (1999). Uniform approximation of discrete-time nonlinear systems. In *IJCNN'99. International Joint Conference on Neural Networks. Proceedings*, volume 1, pages 611–616 vol.1.

[23] Connor, J. A., Walter, D., and McKown, R. (1977). Neural repetitive firing: modifications of the Hodgkin-Huxley axon suggested by experimental results from crustacean axons. *Biophysical Journal*, 18(1):81–102.

[24] Danner, S. M., Hofstoetter, U. S., Freundl, B., Binder, H., Mayr, W., Rattay, F., and Minassian, K. (2015). Human spinal locomotor control is based on flexibly organized burst generators. *Brain*, 138(3):577–588.

[25] Dayan, P. and Abbott, L. F. (2001). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. MIT Press, Cambridge, MA.

[26] Decuyper, J., De Troyer, T., Runacres, M. C., Tiels, K., and Schoukens, J. (2018). On selecting appropriate training data to model an autonomous oscillator. In *International Conference on Noise and Vibration Engineering*, pages 2601–2614, Leuven, Belgium.

[27] Dethier, J., Drion, G., Franci, A., and Sepulchre, R. (2015). A positive feedback at the cellular level promotes robustness and modulation at the circuit level. *Journal of Neurophysiology*.

[28] Drion, G., Franci, A., and Sepulchre, R. (2019). Cellular switches orchestrate rhythmic circuits. *Biological Cybernetics*, 113(1):71–82.

[29] Drion, G., Franci, A., Seutin, V., and Sepulchre, R. (2012). A Novel Phase Portrait for Neuronal Excitability. *PLOS ONE*, 7(8):e41806.

[30] Drion, G., O'Leary, T., Dethier, J., Franci, A., and Sepulchre, R. (2015a). Neuronal behaviors: A control perspective. In *54th IEEE Conference on Decision and Control*, pages 1923–1944.

[31] Drion, G., O'Leary, T., and Marder, E. (2015b). Ion channel degeneracy enables robust and tunable neuronal firing rates. *Proceedings of the National Academy of Sciences*, 112(38):E5361–E5370.

[32] Druckmann, S., Banitt, Y., Gidon, A., Schürmann, F., Markram, H., and Segev, I. (2007). A Novel Multiple Objective Optimization Framework for Constraining Conductance-Based Neuron Models by Experimental Data. *Frontiers in Neuroscience*, 1(1):7–18.

[33] Druckmann, S., Berger, T. K., Schürmann, F., Hill, S., Markram, H., and Segev, I. (2011). Effective Stimuli for Constructing Reliable Neuron Models. *PLOS Computational Biology*, 7(8):e1002133.

[34] Eckmann, J. P. and Ruelle, D. (1985). Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics*, 57(3):617–656.

[35] Enqvist, M. and Ljung, L. (2005). Linear approximations of nonlinear FIR systems for separable input processes. *Automatica*, 41(3):459–473.

[36] Ermentrout, G. B., Galán, R. F., and Urban, N. N. (2008). Reliability, synchrony and noise. *Trends in Neurosciences*, 31(8):428–434.

[37] Ermentrout, G. B. and Terman, D. H. (2010). *Mathematical Foundations of Neuroscience*. Springer, New York.

[38] FitzHugh, R. (1961). Impulses and Physiological States in Theoretical Models of Nerve Membrane. *Biophysical Journal*, 1(6):445–466.

[39] Forni, F. and Sepulchre, R. (2014). Differential analysis of nonlinear systems: Revisiting the pendulum example. In *53rd IEEE Conference on Decision and Control*, pages 3848–3859.

[40] Forssell, U. and Ljung, L. (1999). Closed-loop identification revisited. *Automatica*, 35(7):1215–1241.

[41] Franci, A., Drion, G., and Sepulchre, R. (2012). An Organizing Center in a Planar Model of Neuronal Excitability. *SIAM Journal on Applied Dynamical Systems*, 11(4):1698–1722.

[42] Franci, A., Drion, G., and Sepulchre, R. (2017). Robust and tunable bursting requires slow positive feedback. *Journal of Neurophysiology*, 119(3):1222–1234.

[43] Fröhlich, F. (2016). *Network Neuroscience*. Academic Press, Cambridge, MA.

[44] Galán, R. F., Bard Ermentrout, G., and Urban, N. N. (2007). Reliability and stochastic synchronization in type I vs. type II neural oscillators. *Neurocomputing*, 70(10):2102–2106.

[45] Geit, W. V., Schutter, E. D., and Achard, P. (2008). Automated neuron model optimization techniques: a review. *Biological Cybernetics*, 99(4-5):241–251.

[46] Gerstner, W., Kistler, W. M., Naud, R., and Paninski, L. (2014). *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge University Press, Cambridge, UK.

[47] Gerstner, W. and Naud, R. (2009). How Good Are Neuron Models? *Science*, 326(5951):379–380.

[48] Gevers, M. (1993). Towards a Joint Design of Identification and Control ? In *Essays on Control*, Progress in Systems and Control Theory, pages 111–151. Birkhäuser, Boston, MA.

[49] Goldwyn, J. H. and Shea-Brown, E. (2011). The What and Where of Adding Channel Noise to the Hodgkin-Huxley Equations. *PLOS Computational Biology*, 7(11):e1002247.

[50] Gouwens, N. (2017). Biophysical Modeling - perisomatic. Technical report.

[51] Grashow, R., Brookings, T., and Marder, E. (2009). Reliable neuromodulation from circuits with variable underlying structure. *Proceedings of the National Academy of Sciences*, 106(28):11742–11746.

[52] Grashow, R., Brookings, T., and Marder, E. (2010). Compensation for Variable Intrinsic Neuronal Excitability by Circuit-Synaptic Interactions. *Journal of Neuroscience*, 30(27):9145–9156.

[53] Guertin, P. A. (2009). The mammalian central pattern generator for locomotion. *Brain Research Reviews*, 62(1):45–56.

[54] Heuberger, P. S. C., Hof, P. M. J. V. d., and Bosgra, O. H. (1995). A generalized orthonormal basis for linear dynamical systems. *IEEE Transactions on Automatic Control*, 40(3):451–465.

[55] Heuberger, P. S. C., Hof, P. M. J. v. d., and Wahlberg, B. (2005). *Modelling and Identification with Rational Orthogonal Basis Functions*. Springer, London.

[56] Hille, B. (1984). *Ionic channels of excitable membranes*. Sinauer Associates, Sunderland, MA.

[57] Hodgkin, A. L. and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4):500–544.

[58] Horn, R. A. and Johnson, C. R. (2012). *Matrix Analysis*. Cambridge University Press, Cambridge, UK, 2nd edition.

[59] Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366.

[60] Huys, Q. J. M., Ahrens, M. B., and Paninski, L. (2006). Efficient Estimation of Detailed Single-Neuron Models. *Journal of Neurophysiology*, 96(2):872–890.

[61] Isidori, A. (1995). *Nonlinear Control Systems*. Communications and Control Engineering. Springer-Verlag, London, 3 edition.

[62] Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, 14(6):1569–1572.

[63] Izhikevich, E. M. (2007). *Dynamical Systems in Neuroscience*. MIT Press, Cambridge, MA.

[64] Jensen, R. V. (1998). Synchronization of randomly driven nonlinear oscillators. *Physical Review E*, 58(6):R6907–R6910.

[65] Jensen, R. V. (2002). Synchronization of driven nonlinear oscillators. *American Journal of Physics*, 70(6):607–619.

[66] Keener, J., Sneyd, J., Antman, S. S., Marsden, J. E., and Sirovich, L., editors (2009). *Mathematical Physiology*, volume 8/1 of *Interdisciplinary Applied Mathematics*. Springer, New York, NY.

[67] Khalil, H. K. (2002). *Nonlinear Systems*. Prentice Hall, Upper Saddle River, NJ, 3 edition.

[68] Koch, C. (2004). *Biophysics of Computation: Information Processing in Single Neurons*. Oxford University Press, Oxford, UK.

[69] Kreyszig, E. (1989). *Introductory Functional Analysis with Applications*. John Wiley & Sons, Hoboken, NJ.

[70] Lapicque, L. (1907). Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarisation. *J. Physiol. Pathol. Gen.*, 9:620–635.

[71] Lepora, N. F., Overton, P. G., and Gurney, K. (2012). Efficient fitting of conductance-based model neurons from somatic current clamp. *Journal of Computational Neuroscience*, 32(1):1–24.

[72] Lewi, J., Butera, R., and Paninski, L. (2009). Sequential Optimal Design of Neurophysiology Experiments. *Neural Computation*, 21(3):619–687.

[73] Lewi, J., Schneider, D. M., Woolley, S. M. N., and Paninski, L. (2011). Automating the design of informative sequences of sensory stimuli. *Journal of Computational Neuroscience*, 30(1):181–200.

[74] Linaro, D. and Giugliano, M. (2014). Markov Models of Ion Channels. In Jaeger, D. and Jung, R., editors, *Encyclopedia of Computational Neuroscience*, pages 1–14. Springer, New York, NY.

[75] Liu, Z., Golowasch, J., Marder, E., and Abbott, L. F. (1998). A Model Neuron with Activity-Dependent Conductances Regulated by Multiple Calcium Sensors. *Journal of Neuroscience*, 18(7):2309–2320.

[76] Ljung, L. (1978). Convergence analysis of parametric identification methods. *IEEE Transactions on Automatic Control*, 23(5):770–783.

[77] Ljung, L. (1999). *System Identification: Theory for the User*. Prentice Hall PTR, Upper Saddle River, NJ.

[78] Ljung, L. (2001). Estimating Linear Time-invariant Models of Nonlinear Time-varying Systems. *European Journal of Control*, 7(2):203–219.

[79] Ljung, L. (2009). Experiments with Identification of Continuous Time Models. *IFAC Proceedings Volumes*, 42(10):1175–1180.

[80] Ljung, L. (2010). Perspectives on system identification. *Annual Reviews in Control*, 34(1):1–12.

[81] Lohmiller, W. and Slotine, J.-J. E. (1998). On Contraction Analysis for Non-linear Systems. *Automatica*, 34(6):683–696.

[82] MacKay, D. J. C. (1992). Information-Based Objective Functions for Active Data Selection. *Neural Computation*, 4(4):590–604.

[83] Mäkilä, P. M. and Partington, J. R. (2003). On linear models for nonlinear systems. *Automatica*, 39(1):1–13.

[84] Mäkilä, P. M. and Partington, J. R. (2004). Least-squares LTI approximation of nonlinear systems and quasistationarity analysis. *Automatica*, 40(7):1157–1169.

[85] Manchester, I. R., Tobenkin, M. M., and Wang, J. (2011). Identification of nonlinear systems with stable oscillations. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 5792–5797.

[86] Marder, E. and Bucher, D. (2007). Understanding Circuit Dynamics Using the Stomatogastric Nervous System of Lobsters and Crabs. *Annual Review of Physiology*, 69(1):291–316.

[87] Marder, E., O'Leary, T., and Shruti, S. (2014). Neuromodulation of Circuits with Variable Parameters: Single Neurons and Small Circuits Reveal Principles of State-Dependent and Robust Neuromodulation. *Annual Review of Neuroscience*, 37(1):329–346.

[88] Markovsky, I. and Van Huffel, S. (2007). Overview of total least-squares methods. *Signal Processing*, 87(10):2283–2302.

[89] Matsumoto, T. (1984). A chaotic attractor from Chua's circuit. *IEEE Transactions on Circuits and Systems*, 31(12):1055–1058.

[90] McDougal, R. A., Morse, T. M., Carnevale, T., Marenco, L., Wang, R., Migliore, M., Miller, P. L., Shepherd, G. M., and Hines, M. L. (2017). Twenty years of ModelDB and beyond: building essential modeling tools for the future of neuroscience. *Journal of Computational Neuroscience*, 42(1):1–10.

[91] Meliza, C. D., Kostuk, M., Huang, H., Nogaret, A., Margoliash, D., and Abarbanel, H. D. I. (2014). Estimating parameters and predicting membrane voltages with conductance-based neuron models. *Biological Cybernetics*, 108(4):495–516.

[92] Mensi, S., Naud, R., Pozzorini, C., Avermann, M., Petersen, C. C. H., and Gerstner, W. (2012). Parameter extraction and classification of three cortical neuron types reveals two distinct adaptation mechanisms. *Journal of Neurophysiology*, 107(6):1756–1775.

[93] Milescu, L. S., Akk, G., and Sachs, F. (2005). Maximum Likelihood Estimation of Ion Channel Kinetics from Macroscopic Currents. *Biophysical Journal*, 88(4):2494–2515.

[94] Nagumo, J., Arimoto, S., and Yoshizawa, S. (1962). An Active Pulse Transmission Line Simulating Nerve Axon. *Proceedings of the IRE*, 50(10):2061–2070.

[95] Nicolas-Alonso, L. F. and Gomez-Gil, J. (2012). Brain Computer Interfaces, a Review. *Sensors*, 12(2):1211–1279.

[96] Ninness, B. and Gustafsson, F. (1997). A unifying construction of orthonormal bases for system identification. *IEEE Transactions on Automatic Control*, 42(4):515–521.

[97] Nogaret, A., Meliza, C. D., Margoliash, D., and Abarbanel, H. D. I. (2016). Automatic Construction of Predictive Neuron Models through Large Scale Assimilation of Electrophysiological Data. *Scientific Reports*, 6:32749.

[98] Novara, C., Vincent, T., Hsu, K., Milanese, M., and Poolla, K. (2011). Parametric identification of structured nonlinear systems. *Automatica*, 47(4):711–721.

[99] O'Leary, T., Sutton, A. C., and Marder, E. (2015). Computational models in the age of large datasets. *Current Opinion in Neurobiology*, 32:87–94.

[100] Paduart, J., Lauwers, L., Swevers, J., Smolders, K., Schoukens, J., and Pintelon, R. (2010). Identification of nonlinear systems using Polynomial Nonlinear State Space models. *Automatica*, 46(4):647–656.

[101] Paninski, L. (2004). Maximum likelihood estimation of cascade point-process neural encoding models. *Network: Computation in Neural Systems*, 15(4):243–262.

[102] Paninski, L. (2005). Asymptotic Theory of Information-Theoretic Experimental Design. *Neural Computation*, 17(7):1480–1507.

[103] Paninski, L., Ahmadian, Y., Ferreira, D. G., Koyama, S., Rahnama Rad, K., Vidne, M., Vogelstein, J., and Wu, W. (2010). A new look at state-space models for neural data. *Journal of Computational Neuroscience*, 29(1):107–126.

[104] Paninski, L., Pillow, J., and Lewi, J. (2007). Statistical models for neural encoding, decoding, and optimal stimulus design. In Cisek, P., Drew, T., and Kalaska, J. F., editors, *Progress in Brain Research*, volume 165 of *Computational Neuroscience: Theoretical Insights into Brain Function*, pages 493–507. Elsevier.

[105] Paninski, L., Pillow, J., and Simoncelli, E. (2005). Comparing integrate-and-fire models estimated using intracellular and extracellular data. *Neurocomputing*, 65:379–385.

[106] Paninski, L., Pillow, J. W., and Simoncelli, E. P. (2004). Maximum Likelihood Estimation of a Stochastic Integrate-and-Fire Neural Encoding Model. *Neural Computation*, 16(12):2533–2561.

[107] Park, J. and Sandberg, I. W. (1992). Criteria for the approximation of nonlinear systems. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 39(8):673–676.

[108] Pecora, L. M. and Carroll, T. L. (1990). Synchronization in chaotic systems. *Physical Review Letters*, 64(8):821–824.

[109] Pecora, L. M. and Carroll, T. L. (2015). Synchronization of chaotic systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(9):097611.

[110] Pike, F. G., Goddard, R. S., Suckling, J. M., Ganter, P., Kasthuri, N., and Paulsen, O. (2000). Distinct frequency preferences of different types of rat hippocampal neurones in response to oscillatory input currents. *The Journal of Physiology*, 529(Pt 1):205–213.

[111] Pikovsky, A. and Politi, A. (2016). *Lyapunov Exponents: A Tool to Explore Complex Dynamics*. Cambridge University Press, Cambridge, UK.

[112] Pikovsky, A. S. (1992). Statistics of trajectory separation in noisy dynamical systems. *Physics Letters A*, 165(1):33–36.

[113] Pillow, J. and Park, M. (2016). Adaptive Bayesian Methods for Closed-Loop Neurophysiology. In *Closed Loop Neuroscience*, pages 3–18. Academic Press, Cambridge, MA.

[114] Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E. J., and Simoncelli, E. P. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, 454(7207):995–999.

[115] Pintelon, R. and Schoukens, J. (2012). *System Identification: A Frequency Domain Approach*. John Wiley & Sons, Hoboken, NJ.

[116] Pogromsky, A. Y. (1998). Passivity Based Design of Synchronizing Systems. *International Journal of Bifurcation and Chaos*, 08(02):295–319.

[117] Prinz, A. A., Abbott, L. F., and Marder, E. (2004). The dynamic clamp comes of age. *Trends in Neurosciences*, 27(4):218–224.

[118] Puil, E., Gimbarzevsky, B., and Miura, R. M. (1986). Quantification of membrane properties of trigeminal root ganglion neurons in guinea pigs. *Journal of Neurophysiology*, 55(5):995–1016.

[119] Sandberg, I. W. (1965). Some results on the theory of physical systems governed by nonlinear functional equations. *The Bell System Technical Journal*, 44(5):871–898.

[120] Sandberg, I. W. (1991). Approximation theorems for discrete-time systems. *IEEE Transactions on Circuits and Systems*, 38(5):564–566.

[121] Sandberg, I. W. (1992). Approximately-finite memory and input-output maps. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 39(7):549–556.

[122] Sandberg, I. W. (1994). Approximately finite memory and the circle criterion. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 41(7):473–476.

[123] Schön, T. B., Lindsten, F., Dahlin, J., Wågberg, J., Naesseth, C. A., Svensson, A., and Dai, L. (2015). Sequential Monte Carlo Methods for System Identification. *IFAC-PapersOnLine*, 48(28):775–786.

[124] Schön, T. B., Wills, A., and Ninness, B. (2011). System identification of nonlinear state-space models. *Automatica*, 47(1):39–49.

[125] Schoukens, J., Pintelon, R., Dobrowiecki, T., and Rolain, Y. (2005). Identification of linear systems with nonlinear distortions. *Automatica*, 41(3):491–504.

[126] Schoukens, J., Relan, R., and Schoukens, M. (2017). Discrete Time Approximation of Continuous Time Nonlinear State Space Models. *IFAC-PapersOnLine*, 50(1):8339–8346.

[127] Schoukens, M., Marconato, A., Pintelon, R., Vandersteen, G., and Rolain, Y. (2015). Parametric identification of parallel Wiener–Hammerstein systems. *Automatica*, 51:111–122.

[128] Schoukens, M. and Rolain, Y. (2012). Parametric Identification of Parallel Wiener Systems. *IEEE Transactions on Instrumentation and Measurement*, 61(10):2825–2832.

[129] Schoukens, M. and Tiels, K. (2017). Identification of block-oriented nonlinear systems starting from linear approximations: A survey. *Automatica*, 85:272–292.

[130] Schram, G., Verhaegen, M., and Krijgsman, A. (1996). System Identification with Orthogonal Basis Functions and Neural Networks. *IFAC Proceedings Volumes*, 29(1):4150–4155.

[131] Schrama, R. (1992). *Approximate Identification and Control Design*. PhD thesis, Delft University of Technology.

[132] Sepulchre, R. (2006). Oscillators as Systems and Synchrony as a Design Principle. In Menini, L., Zaccarian, L., and Abdallah, C. T., editors, *Current Trends in Nonlinear Systems and Control*, Systems and Control: Foundations & Applications, pages 123–141. Birkhäuser, Boston, MA.

[133] Sepulchre, R., Drion, G., and Franci, A. (2018). Excitable Behaviors. In *Emerging Applications of Control and Systems Theory*, pages 269–280. Springer International Publishing.

[134] Sepulchre, R., Jankovic, M., and Kokotovic, P. V. (1997). *Constructive Nonlinear Control*. Communications and Control Engineering. Springer-Verlag, London.

[135] Sharp, A. A., Skinner, F. K., and Marder, E. (1996). Mechanisms of oscillation in dynamic clamp constructed two-cell half-center circuits. *Journal of Neurophysiology*, 76(2):867–883.

[136] Skelton, R. E. (1989). Model error concepts in control design. *International Journal of Control*, 49(5):1725–1753.

[137] Slotine, J.-J. E. (2003). Modular stability tools for distributed computation and control. *International Journal of Adaptive Control and Signal Processing*, 17(6):397–416.

[138] Sontag, E. D. (1993). Some Topics in Neural Networks and Control. Technical Report LS93-02, Department of Mathematics, Rutgers University.

[139] Stan, G. B. and Sepulchre, R. (2007). Analysis of Interconnected Oscillators by Dissipativity Theory. *IEEE Transactions on Automatic Control*, 52(2):256–270.

[140] Strogatz, S. H. (2014). *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. Westview Press, Boulder, CO, 2 edition.

[141] Teramae, J.-n. and Tanaka, D. (2004). Robustness of the noise-induced phase synchronization in a general class of limit cycle oscillators. *Physical review letters*, 93(20):204103.

[142] Tiels, K. and Schoukens, J. (2014). Wiener system identification with generalized orthonormal basis functions. *Automatica*, 50(12):3147–3154.

[143] Toth, B. A., Kostuk, M., Meliza, C. D., Margoliash, D., and Abarbanel, H. D. I. (2011). Dynamical estimation of neuron and network properties I: variational methods. *Biological Cybernetics*, 105(3):217–237.

[144] Truccolo, W., Eden, U. T., Fellows, M. R., Donoghue, J. P., and Brown, E. N. (2005). A Point Process Framework for Relating Neural Spiking Activity to Spiking History, Neural Ensemble, and Extrinsic Covariate Effects. *Journal of Neurophysiology*, 93(2):1074–1089.

[145] Ujfalussy, B. B., Makara, J. K., Lengyel, M., and Branco, T. (2018). Global and Multiplexed Dendritic Computations under In Vivo-like Conditions. *Neuron*, 100(3):579–592.e5.

[146] Van den Hof, P. (1998). Closed-loop issues in system identification. *Annual Reviews in Control*, 22:173–186.

[147] Van Den Hof, P. M. J. and Schrama, R. J. P. (1995). Identification and control — Closed-loop issues. *Automatica*, 31(12):1751–1770.

[148] Van Pottelbergh, T., Drion, G., and Sepulchre, R. (2018). Robust Modulation of Integrate-and-Fire Models. *Neural Computation*, 30(4):987–1011.

[149] Wang, W. and Slotine, J.-J. E. (2004). On partial contraction analysis for coupled nonlinear oscillators. *Biological Cybernetics*, 92(1):38–53.

[150] Willems, J. C. (2007). The Behavioral Approach to Open and Interconnected Systems. *IEEE Control Systems Magazine*, 27(6):46–99.

[151] Wolf, A., Swift, J. B., Swinney, H. L., and Vastano, J. A. (1985). Determining Lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena*, 16(3):285–317.

[152] Yuz, J. I. and Goodwin, G. C. (2005). On sampled-data models for nonlinear systems. *IEEE Transactions on Automatic Control*, 50(10):1477–1489.

[153] Zhou, C. and Kurths, J. (2002). Noise-Induced Phase Synchronization and Synchronization Transitions in Chaotic Oscillators. *Physical Review Letters*, 88(23):230602.

[154] Zhou, C. and Kurths, J. (2003). Noise-induced synchronization and coherence resonance of a Hodgkin–Huxley model of thermally sensitive neurons. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 13(1):401–409.

[155] Zhou, K., Doyle, J. C., and Glover, K. (1996). *Robust and Optimal Control*. Prentice Hall, Upper Saddle River, NJ.